MOLECULAR BIOLOGY

*An International Series of Monographs and Textbooks*

[illegible] BIOLOGICAL [illegible]

MBL/WHOI
0 0301 0011707 3

# Molecular Genetics

## Part I

# *Molecular Biology*

*An International Series of Monographs and Textbooks*

*Edited by*

NATHAN O. KAPLAN
*Graduate Department of Biochemistry*
*Brandeis University*
*Waltham, Massachusetts*

HAROLD A. SCHERAGA
*Department of Chemistry*
*Cornell University*
*Ithaca, New York*

*Volume 1.* HAROLD A. SCHERAGA. Protein Structure. 1961

*Volume 2.* STUART A. RICE AND MITSURU NAGASAWA. Polyelectrolyte Solutions: A Theoretical Introduction. *With a contribution by Herbert Morawetz.* 1961

*Volume 3.* SIDNEY UDENFRIEND. Fluorescence Assay in Biology and Medicine. 1962

*Volume 4.* J. HERBERT TAYLOR (Editor). Molecular Genetics. (In two parts.) Part I—1963. Part II in preparation.

*Other Volumes in Preparation*

QH
431
T28
c.2

# Molecular Genetics

*Edited by*

J. HERBERT TAYLOR
*Departments of Botany and Zoology*
*Columbia University*
*New York, New York*

## Part I

1963

ACADEMIC PRESS · NEW YORK AND LONDON

Copyright © 1963, by Academic Press Inc.

ALL RIGHTS RESERVED

NO PART OF THIS BOOK MAY BE REPRODUCED IN ANY FORM
BY PHOTOSTAT, MICROFILM, OR ANY OTHER MEANS,
WITHOUT WRITTEN PERMISSION FROM THE PUBLISHERS.

ACADEMIC PRESS INC.
111 Fifth Avenue
New York 3, N. Y.

*United Kingdom Edition*
Published by
ACADEMIC PRESS INC. (London) Ltd.
Berkeley Square House
Berkeley Square, London, W.1.

*Library of Congress Catalog Card Number 62-2214*

PRINTED IN THE UNITED STATES OF AMERICA

# Contributors to Part I

CORRADO BAGLIONI,* *Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts*

MAURICE J. BESSMAN, *McCollum-Pratt Institute, The Johns Hopkins University, Baltimore, Maryland*

JOHN BISHOP, *Department of Biochemistry, University of Kentucky College of Medicine, Lexington, Kentucky*

R. J. BRITTEN, *Carnegie Institution of Washington, Department of Terrestrial Magnetism, Washington, D.C.*

H. FRAENKEL-CONRAT, *Virus Laboratory, University of California, Berkeley, California*

ERNST FREESE,† *Department of Genetics, University of Wisconsin, Madison, Wisconsin*

KARL G. LARK, *Department of Microbiology, St. Louis University School of Medicine, St. Louis, Missouri*

B. J. MCCARTHY, *Carnegie Institution of Washington, Department of Terrestrial Magnetism, Washington, D.C.*

R. B. ROBERTS, *Carnegie Institution of Washington, Department of Terrestrial Magnetism, Washington, D.C.*

RICHARD SCHWEET, *Department of Biochemistry, University of Kentucky College of Medicine, Lexington, Kentucky*

J. HERBERT TAYLOR, *Department of Botany and Department of Zoology, Columbia University, New York, New York*

CHARLES A. THOMAS, JR., *The Johns Hopkins University, Baltimore, Maryland*

A. TSUGITA,‡ *Virus Laboratory, University of California, Berkeley, California*

ELLIOT VOLKIN, *Biology Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee*

---

* Present address: International Laboratory for Genetics and Biophysics, Naples, Italy.

† Present address: Laboratory of Molecular Biology, National Institutes of Health—NINDB, Bethesda, Maryland.

‡ Present address: Research Institute for Protein, Osaka University, Osaka, Japan.

# Preface

Although genetics is one of the youngest subdivisions of biology, it has come to occupy a central position and to serve as a focal point not only for the interests of biologists, but also for diverse groups who use designations such as biochemists, physical chemists, biophysicists, virologists, geneticists, and cytologists or cell biologists. A consideration of the groups involved indicates that the emphasis in this work has shifted from the classical approaches of genetics to a molecular view of the mechanism of heredity; therefore "Molecular Genetics" is an appropriate title. Many of the problems are being considered from the molecular point of view even though the structures and mechanisms involved cannot yet be precisely defined in molecular terms.

In a rapidly developing field of science, the ability to keep informed of progress becomes a major problem. When the groups involved have such diverse backgrounds, interests, and locations among the various established disciplines, the task becomes even more difficult. Fortunately there is no dearth of scientific meetings and special symposia which consider various aspects of the ill-defined area to be included under molecular genetics. Many of these meetings and discussions result in publications. However, a need was felt for a more consistent and thorough coverage, which would highlight some of the interesting and, hopefully, significant advances, provide a useful reference source, and allow our interested colleagues, who work in these or related areas, to bring themselves up to date. The material should also be useful for teaching and training at the graduate and upper undergraduate levels. Only time can render a verdict on the success of these aims.

As indicated by the designation of this as Part I, several important topics in molecular genetics have been missed or touched lightly. Although plans are still indefinite we hope to fill the gaps in a second and possibly third volume. Some of the most exciting concepts are still not sufficiently crystallized to warrant recording even in an advanced work such as this. As some degree of stabilization occurs, and willing and able writers are found, we trust the gaps will be closed. Nevertheless, a unity and narrative form were kept in mind in planning the volume. The first five chapters are concerned with DNA—its replication and organization into chromosomes, control mechanisms in its replication, and its mutability. The next two chapters deal with the biosynthesis of RNA, its role in translating and transferring the genetic information, and the sites of its synthesis, organization, and function. The last three chapters deal with the genetic aspects of protein synthesis and the transcription of the genetic code into functional entities. A few of the areas not covered will be obvious; for example, exciting conceptual developments on control mechanism in the transcription of the genetic code are touched lightly. The cell-free synthesis of RNA, which was advancing so rapidly as this volume was written, is yet to be covered in detail. Another topic, genetic recombination, which is of deep concern and great significance to geneticists, has been largely ignored, perhaps with good reason. A major conceptual revolution, with some indication that the topic can be treated in molecular terms, is now in the making.

The chapters were not intended as comprehensive reviews, but represent personal accounts of developments best known to each writer, who in all cases was involved in making significant advances in the respective areas. However, there was an attempt to present a balanced view, to cite related findings and most of all to write a readable and up-to-date account. In a field which moves as fast as molecular genetics, publications tend to be outdated before they are off the press. To circumvent this, in part at least, we hope we have covered some material of lasting importance, presented the newest developments known to us while writing, and in some cases looked into the future with a bit of speculation that will hopefully prove stimulating and productive.

The Editor is indebted to the contributors who paused in their research long enough to make their contribution and who responded to the stated objective of finishing the manuscripts together so that all of the chapters would have the same degree of timeliness, except for the limitations imposed by the nature of the material covered. For expediting the work and for skillfully managing many of the technical details, the publishers are to be given much of the credit. To all of those who contributed by supplying unpublished figures and data or made material in press available to the various authors, special credit is due. Many of them are acknowledged in the various chapters, but for contributions of unpublished material to Chapter II, I wish to thank Joseph G. Gall of the University of Minnesota, George D. Pappas and Philip W. Brandt of Columbia University, and Oscar L. Miller, Biology Division, Oak Ridge National Laboratory.

J. HERBERT TAYLOR

# Contents of Part I

CONTRIBUTORS TO PART I . . . . . . . . . . . . . . v

PREFACE . . . . . . . . . . . . . . . . . . . vii

## *Chapter* I

## The Replication of DNA in Cell-Free Systems

Maurice J. Bessman

I. Introduction . . . . . . . . . . . . . . . . . 1
II. Synthesis of Deoxyribonucleotides . . . . . . . . . . 3
III. Phosphorylation of Deoxyribonucleotides . . . . . . . . . 12
IV. Synthesis of DNA by *Escherichia coli* Polymerase . . . . . . 17
V. Synthesis of DNA by Calf Thymus Polymerase . . . . . . . 50
VI. Synthesis of DNA in Bacteriophage-Infected *Escherichia coli* . . . 54
VII. Conclusions . . . . . . . . . . . . . . . . . 60
Acknowledgments . . . . . . . . . . . . . . . . 60
References . . . . . . . . . . . . . . . . . . 60

## *Chapter* II

## The Replication and Organization of DNA in Chromosomes

J. Herbert Taylor

I. Introduction . . . . . . . . . . . . . . . . . 65
II. A Morphological Picture of the Chromosome . . . . . . . . 67
III. Replication of DNA . . . . . . . . . . . . . . . 72
IV. Exchanges between DNA Subunits and Genetic Recombination . . 82
V. Sequences of Replication of DNA in Chromosomes . . . . . . 88
VI. Chromosome Reproduction in Relation to DNA Replication and Protein Synthesis . . . . . . . . . . . . . . . 89
VII. Molecular Models for Organization of DNA into Chromosomes . . 96
References . . . . . . . . . . . . . . . . . . 108

## *Chapter* III

## The Organization of DNA in Bacteriophage and Bacteria

CHARLES A. THOMAS, JR.

I. Bacteriophage DNA . . . . . . . . . . . . . . 113
II. Bacterial DNA . . . . . . . . . . . . . . 141
Acknowledgments . . . . . . . . . . . . . . 148
References . . . . . . . . . . . . . . . 148

## *Chapter* IV

## Cellular Control of DNA Biosynthesis

KARL G. LARK

I. Introduction . . . . . . . . . . . . . . . 153
II. The Role of DNA Precursors in Controlling DNA Synthesis . . . 162
III. Role of Protein or RNA Synthesis in Controlling DNA Synthesis . . 173
IV. The Physical State of DNA as a Factor Controlling its Replication . 180
V. Other Factors Affecting DNA Biosynthesis and the DNA Cycle . . 190
VI. Summary, Conclusion, and Opinions . . . . . . . . . 197
Acknowledgments . . . . . . . . . . . . . . 200
References . . . . . . . . . . . . . . . 200

## *Chapter* V

## Molecular Mechanism of Mutations

ERNST FREESE

I. Introduction . . . . . . . . . . . . . . . 207
II. Structure, Duplication, and Alteration of the Hereditary Material . . 215
III. Biochemistry of Mutagens and Theory of Base Alterations . . . 217
IV. The Extent of Point Mutations . . . . . . . . . . 248
V. Mutagenic Specificity . . . . . . . . . . . . 252
References . . . . . . . . . . . . . . . 264

## *Chapter* VI

## Biosynthesis of RNA in Relation to Genetic Coding Problems

ELLIOT VOLKIN

I. Introduction . . . . . . . . . . . . . . . . . . 271
II. Evidence Implicating an RNA Species as a Carrier of Genetic Information . . . . . . . . . . . . . . . . . . 273
III. Mechanism of Messenger RNA Function Based on *in Vivo* Experiments . . . . . . . . . . . . . . . . . . 277
IV. Attempts at Isolation of a Messenger RNA . . . . . . . . . 279
V. *In Vitro* Synthesis of RNA . . . . . . . . . . . . . . 280
VI. RNA Involvement in *in Vitro* Protein Synthesis . . . . . . . 281
VII. The Triplet Code . . . . . . . . . . . . . . . . 284
VIII. Summary . . . . . . . . . . . . . . . . . . 286
References . . . . . . . . . . . . . . . . . . 286

## *Chapter* VII

## Kinetic Studies of the Synthesis of RNA and Ribosomes

R. B. ROBERTS, R. J. BRITTEN, AND B. J. MCCARTHY

I. Introduction . . . . . . . . . . . . . . . . . . 292
II. Properties of Ribosomes *in Vitro* . . . . . . . . . . . 299
III. Fractionation of Ribosomes . . . . . . . . . . . . . 308
IV. Metabolic Pools . . . . . . . . . . . . . . . . . 311
V. Kinetics of Incorporation into RNA of Ribosomes . . . . . . 315
VI. Kinetic Studies of Phenol-Extracted RNA . . . . . . . . . 329
VII. Kinetics of Incorporation into Protein . . . . . . . . . . 331
VIII. Composition of Newly Formed RNA . . . . . . . . . . 335
IX. Discussion . . . . . . . . . . . . . . . . . . 342
X. Epilogue . . . . . . . . . . . . . . . . . . 350
References . . . . . . . . . . . . . . . . . . 350

## *Chapter* VIII

## Protein Synthesis in Relation to Gene Action

RICHARD SCHWEET AND JOHN BISHOP

I. Introduction . . . . . . . . . . . . . . . . . . 353
II. Formation of Amino Acyl-RNA Compounds . . . . . . . . 355

III. Steps Requiring Participation of the Ribosome . . . . . . . 362
IV. Information Transfer in Protein Synthesis . . . . . . . . . 380
References . . . . . . . . . . . . . . . . 400

## *Chapter* IX

## Correlations between Genetics and Chemistry of Human Hemoglobins

CORRADO BAGLIONI

I. Introduction . . . . . . . . . . . . . . . 405
II. The Separation of Human Hemoglobins by Chromatography and Electrophoresis . . . . . . . . . . . . 406
III. Normal Hemoglobins . . . . . . . . . . . . 409
IV. Abnormal Hemoglobins . . . . . . . . . . . . 421
V. The Control of the Synthesis of Hemoglobin . . . . . . . 456
Acknowledgments . . . . . . . . . . . . . 469
References . . . . . . . . . . . . . . . . 469

## *Chapter* X

## Contributions from TMV Studies to the Problem of Genetic Information Transfer and Coding

A. TSUGITA AND H. FRAENKEL-CONRAT

I. Introduction . . . . . . . . . . . . . . . 477
II. TMV-Protein . . . . . . . . . . . . . . 479
III. TMV-RNA . . . . . . . . . . . . . . . 487
IV. Viral Infectivity . . . . . . . . . . . . . 489
V. Natural Strains of TMV . . . . . . . . . . . 491
VI. Chemically Produced Strains . . . . . . . . . . 497
VII. The Coding Problem . . . . . . . . . . . . 509
References . . . . . . . . . . . . . . . . 517

AUTHOR INDEX . . . . . . . . . . . . . . . 521

SUBJECT INDEX . . . . . . . . . . . . . . . 540

*Chapter* I

# The Replication of DNA in Cell-Free Systems[1]

MAURICE J. BESSMAN

*McCollum-Pratt Institute,*
*The Johns Hopkins University,*
*Baltimore, Maryland*

I. Introduction . . . . . . . . . . . . . . . . . 1
II. Synthesis of Deoxyribonucleotides . . . . . . . . . . . 3
  A. Origin of the Deoxyribosides . . . . . . . . . . . 3
  B. Synthesis of Thymidylic Acid . . . . . . . . . . . . 6
  C. Synthesis of 5-Hydroxymethyldeoxycytidylic Acid (dHMP) . . 10
III. Phosphorylation of Deoxyribonucleotides . . . . . . . . 12
  A. Deoxynucleoside Monophosphate Kinases . . . . . . . . 12
  B. Deoxynucleoside Diphosphate Kinases . . . . . . . . . 16
IV. Synthesis of DNA by *Escherichia coli* Polymerase . . . . . . 17
  A. Measurement of DNA Synthesis . . . . . . . . . . . 17
  B. Requirements of the Reaction . . . . . . . . . . . 20
  C. The Limited Reaction . . . . . . . . . . . . . 23
  D. Nature of the Primer . . . . . . . . . . . . . 25
  E. Nature of the Product . . . . . . . . . . . . . 27
  F. Primer-Product Relationships . . . . . . . . . . . 29
  G. Synthesis of Deoxyadenylate-Deoxythymidylate and Deoxyguanylate-Deoxycytidylate Polymers . . . . . . . . . . . 40
  H. Substitution of Analogs . . . . . . . . . . . . 46
V. Synthesis of DNA by Calf Thymus Polymerase . . . . . . . 50
VI. Synthesis of DNA in Bacteriophage-Infected *Escherichia coli* . . . 54
  A. Bacteriophage-Specific Polymerase . . . . . . . . . 54
  B. Glucosylation of 5-Hydroxymethylcytosine (HMC) . . . . . 56
VII. Conclusions . . . . . . . . . . . . . . . . . 60
  Acknowledgments . . . . . . . . . . . . . . . . 60
  References . . . . . . . . . . . . . . . . . 60

## I. INTRODUCTION[2]

Most organisms can synthesize all the constituents of the DNA macromolecule from small molecular weight precursors which are also the raw materials for the synthesis of many other cellular constituents. It is probably the synthesis of deoxyribose or a derivative of it which is the point of diversion from this non-specific pool of precursors to the specific sequence of reactions culminating in the polymerization of the deoxyribonucleoside triphosphates to form the DNA molecule. Since several of the later chapters refer to specific reactions in this metabolic chain, the current concepts of the pathways of enzymatic synthesis of deoxyribonucleotides and their phosphorylated derivatives will be discussed briefly with the intention of providing an immediate reference for the ensuing chapters. Figure 1 shows several of the reactions suggested as intermediate steps leading to the synthesis of DNA.

---

[1] Contribution No. 371 of the McCollum-Pratt Institute, The Johns Hopkins University, Baltimore 18, Maryland.

[2] The following abbreviations have been used in this chapter: RNA, DNA: ribonucleic acid, deoxyribonucleic acid; RNase, DNase: ribonuclease, deoxyribo-

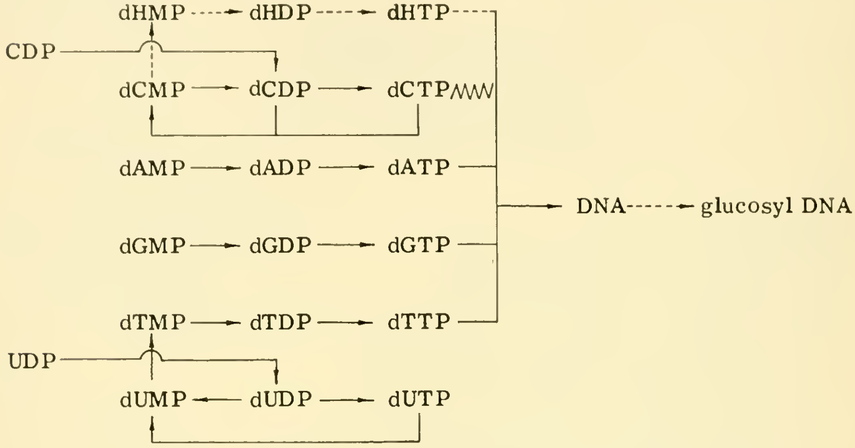


FIG. 1. Enzymatic reactions in *E. coli* leading to the synthesis of DNA. ----, specific to cells infected with T-even phages. ⁓⁓, blocked in cells infected with T-even phages.

In the cell-free system, the replication of DNA implies the net synthesis of exact replicas of the "primer" molecules. For this reason, only those systems sufficiently purified to demonstrate both net synthesis

---

nuclease; AMP, ADP, ATP: adenosine monophosphate, adenosine diphosphate, adenosine triphosphate; CMP, CDP, CTP: cytidine monophosphate, etc.; GMP, GDP, GTP: guanosine monophosphate, etc.; HMP, HDP, HTP: hydroxymethylcytidine monophosphate, etc.; IMP, IDP, ITP: inosine monophosphate, etc.; TMP, TDP, TTP: thymidine monophosphate, etc.; UMP, UDP, UTP: uridine monophosphate, etc.; dAMP, dADP, dATP: deoxyadenosine monophosphate, etc.; DPN, DPNH: diphosphopyridine and its reduced form; TPN, TPNH: triphosphopyridine and its reduced form; FAD, $FADH_2$: flavin adenine dinucleotide and its reduced form; THFA: tetrahydrofolic acid; $P_i$, $PP_i$: inorganic orthophosphate and pyrophosphate; tris: tris(hydroxymethyl)aminomethane.

and a requirement for primer DNA will be considered in this chapter. The influence of primer on the synthetic reaction is a necessary but certainly not a sufficient condition for replication since, as will be discussed later, there are some situations in which the synthetic DNA reflects the composition and structure of the added primer, and other primer-dependent reactions in which the product is clearly different from the primer.

## II. SYNTHESIS OF DEOXYRIBONUCLEOTIDES

### A. ORIGIN OF THE DEOXYRIBOSIDES

Recent experiments with cell-free systems support the earlier indications from tracer experiments in the whole animal that deoxyribosyl compounds are synthesized by a direct reduction of some ribosyl derivatives. Hammarsten *et al.* (1950) suggested that the $N^{15}$-labeled deoxycytidine they isolated from rat liver DNA arose by a direct conversion of the injected $N^{15}$-cytidine without cleavage of the ribosidic linkage, since it was known from the results of several laboratories that free cytosine is not a good precursor of DNA cytosine. Experimental evidence in support of this contention was furnished by the work of Rose and Schweigert (1953), who injected $C^{14}$-cytidine, labeled in both the base and the ribose moieties, into rats and found that the ratio of counts in the base and deoxyribose in the DNA deoxycytidine was the same as the ratio of counts in the base and ribose of the injected nucleoside. Either cytidine was converted to deoxycytidine without cleavage of the ribosidic linkage, or cleavage did occur and the pool sizes of the free cytosine and ribose in the rat were exactly the same. The latter possibility seems too remote to have been seriously entertained. These results were confirmed by Roll *et al.* (1956) and Reichard (1957) and extended to include chick embryo (Reichard, 1958, 1959), Ehrlich ascites cells (Edmunds, 1958) and *Neurospora* (McNutt, 1958a). Cell-free extracts of chick embryo (Reichard, 1959, 1960), *Escherichia coli* (Reichard and Rutberg, 1960; Reichard *et al.*, 1961), and Novikoff hepatoma (Moore and Hurlbert, 1960) were shown to carry out the reduction of ribosyl compounds to deoxyribosyl derivatives and, as is usually the case, it has been through the use of purified, cell-free systems that insight into the mechanism of the interconversion has become possible. Reichard *et al.* (1961) had suggested from isotope dilution experiments and kinetic data that the conversion of cytidine to deoxycytidine nucleotides occurred at the diphosphate level and required ATP, $Mg^{++}$, and TPN. Bertani *et al.* (1961), using a fraction purified 100–200-fold from crude extracts of *Escherichia coli*, have presented clear evidence that the reduction occurs at the diphosphate level and show that optimal synthesis requires the participation of ATP, $Mg^{++}$, and reduced lipoic acid. These results are presented in Table I. It is interesting to note that reduced lipoic acid is an essential component of the system and cannot be replaced by TPNH, DPNH, glutathione, ascorbic acid, 2,3-mercaptopropanol, or oxidized lipoic acid. The authors suggest that the effect of lipoic acid may be indirect, mediated through a flavin coenzyme, but they could not replace lipoic acid by FAD or $FADH_2$. The requirement for ATP suggests the participation of an "activated" intermediate during the reductive process, possibly a phosphorylated, pyrophosphorylated, or adenylated derivative at the 2′-position (Bertani *et al.*, 1961). This same enzyme fraction catalyzes the reduction of UDP to deoxy-UDP, again at the diphosphate level (Table II).

TABLE I
REQUIREMENTS FOR dCDP FORMATION WITH A PURIFIED FRACTION FROM *Escherichia coli* [a,b]

| Omission | Addition | Deoxycytidine phosphates formed (mμmoles) |
|---|---|---|
| None | None | 2.75 |
| Tritium-labeled CDP | Tritium-labeled CMP (0.04 μmole) | 0.25 |
| Tritium-labeled CDP | Tritium-labeled CTP (0.04 μmole) | 0.73 |
| ATP | None | 0.06 |
| $MgCl_2$ | None | 0.03 |
| Lipoic acid | None | 0.05 |
| Lipoic acid | TPNH (0.14 μmole) | 0.04 |

[a] From Bertani *et al.* (1961).

[b] The complete reaction mixture contained 0.04 μmole of $CDPH^3$ ($3 \times 10^6$ cpm per μmole), 0.5 μmole of ATP, 1.5 μmoles of $MgCl_2$, 0.2 μmoles of reduced DL-lipoic acid, 2.5 μmoles of tris-HCl buffer, pH 7.85, 0.1 μmole of mercaptoethanol, and 0.11 mg of fraction B in a final volume of 0.1 ml.

The specificity of this enzyme in respect to nucleoside diphosphates has not been described and the question may be asked whether the same enzyme or similar enzymes account for the net formation of purine deoxyribonucleotides. Indications that the purine deoxyribosyl compounds are formed directly from ribosyl derivatives have been reported from experiments in whole cells (Roll *et al.*, 1956; McNutt, 1958b) and in cell extracts (Reichard, 1960). This probably represents the major pathway for the formation of the purine deoxyribonucleotides, but the possibility of a contribution by the transdeoxyribosidase of McNutt

TABLE II

FORMATION OF DEOXYURIDINE PHOSPHATES FROM DIFFERENT URIDINE PHOSPHATES[a,b]

| Substrate | Deoxyuridine phosphates formed (mμmoles) |
|---|---|
| UMP (28 mμmoles) | 0.003 |
| UDP (36 mμmoles) | 0.340 |
| UTP (38 mμmoles) | 0.024 |

[a] From Bertani *et al.* (1961).

[b] The reaction mixture contained the different $C^{14}$-uridine nucleotides ($11 \times 10^6$ cpm per μmole) together with 0.5 μmole of ATP, 1.2 μmoles of $MgCl_2$, 0.34 μmole of reduced lipoic acid, 2.3 μmole of tris-HCl buffer, pH 7.85, 0.1 μmole of mercaptoethanol, and 0.16 mg of fraction B, in a final volume of 0.1 ml.

(1952) should not be overlooked. This latter enzyme catalyzes the following reaction:

$$\text{Pyrimidine deoxyriboside} + \text{purine} \rightleftharpoons \text{purinedeoxyriboside} + \text{pyrimidine}$$

The importance of this reaction in the *de novo* synthesis of purine deoxyribonucleotides remains to be evaluated.

A pathway for the formation of deoxyribose, not involving nucleoside derivatives, was described by Racker (1952), who demonstrated the condensation of acetaldehyde and glyceraldehyde-3-phosphate to form deoxyribose-5-phosphate. This reaction, catalyzed by enzymes from *E. coli* or animal tissues (deoxyriboaldolase) could be coupled to the phosphodeoxyribomutase of Manson and Lampen (1951) and the nucleoside phosphorylase of Friedkin and Kalckar (1950) and Friedkin and Roberts (1954) to form deoxyribonucleosides according to the following scheme:

$$\text{Acetaldehyde} + \text{glyceraldehyde-3-phosphate} \overset{\text{aldolase}}{\rightleftharpoons} \text{deoxyribose-5-phosphate}$$

$$\text{Deoxyribose-5-phosphate} \overset{\text{mutase}}{\rightleftharpoons} \text{deoxyribose-1-phosphate}$$

$$\text{Deoxyribose-1-phosphate} + \text{base} \overset{\text{phosphorylase}}{\rightleftharpoons} \text{deoxyribonucleoside} + \text{phosphate}$$

The equilibrium of the aldolase reaction is very unfavorable toward synthesis, suggesting that the significance of this reaction may lie in the degradation of deoxyribose-5-phosphate.

In analogy to ribonucleotide synthesis in which 5-phosphoribosyl-1-pyrophosphate is the active pentose intermediate (Kornberg *et al.*, 1955; Remy *et al.*, 1955), one could reason that a deoxyribose analog of this compound might participate in deoxyribonucleotide synthesis:

Deoxyribose-5-phosphate + ATP $\xrightarrow{\text{kinase}}$ 5-phosphodeoxyribosyl-1-pyrophosphate + AMP

5-Phosphodeoxyribosyl-1-pyrophosphate + base → deoxynucleotide + pyrophosphate

Although these reactions were looked for (Kornberg, 1957) no evidence for such a pathway was found.

### B. SYNTHESIS OF THYMIDYLIC ACID

Tracer experiments in whole-cell systems indicated that the general pathway for the synthesis of thymidylic acid involved the condensation of an "active" 1-carbon unit, probably a derivative of folic acid, with a derivative of uridine or cytidine (for review see Crosbie, 1960). Friedkin and Kornberg (1957) described a cell-free preparation from *E. coli* which converted deoxyuridylate to thymidine triphosphate in the presence of tetrahydrofolic acid (THFA), serine, and ATP. The properties of the system are described in Table III. The role of

TABLE III

Conversion of Deoxyuridylic Acid to Thymidine Triphosphate[a]

| | Cpm on Norit |
|---|---|
| Experiment 1 | |
| Complete system[b] | 1650 |
| zero time | 24 |
| minus THFA | 47 |
| − serine | 480 |
| − ATP | 51 |
| − extract | 212 |
| − Mg | 30 |
| − pyridoxal phosphate | 1500 |
| Experiment 2 | |
| Complete system | |
| + serine + THFA | 1870 |
| − serine + THFA | 470 |
| Complete system | |
| + serine + hydroxymethyl THFA | 1650 |
| − serine + hydroxymethyl THFA | 1210 |

[a] From Friedkin and Kornberg (1957).

[b] The complete system (325 μliters; pH 7.4) contained 0.011 μmoles of deoxyuridine-5′-phosphate labeled with $P^{32}$ (2500 cpm), 2.3 μmoles of ATP, 5.7 μmoles of $MgCl_2$, 11.4 μmoles of L-serine, 0.1 μmole of THFA or $N^{10}$-hydroxymethyl THFA, 0.5 μmole of cysteine, 2.5 μmoles of inorganic orthophosphate, 0.1 μmole of pyridoxal phosphate, 21 μmoles of tris, thymidylate kinase, and an extract of *E. coli* treated with Dowex-1-formate.

serine can be explained by the observation that the active 1-carbon unit which ultimately becomes the methyl group of thymine can be derived from the $\beta$-carbon of serine (Elwyn and Sprinson, 1954). The requirement for serine could be replaced by synthetic hydroxymethyltetrahydrofolic acid which suggested that this was the active 1-carbon intermediate generated in this system. These observations were extended to animal tissues by Phear and Greenberg (1957) and Humphreys and Greenberg (1958) and to a serineless mutant of *E. coli* by Birnie and Crosbie (1958). The bulk of the evidence implicates deoxyuridine monophosphate as the immediate precursor of thymidylate in these systems. Blakely (1957), however, has suggested from his experiments that deoxyuridine rather than the corresponding nucleotide is the acceptor of the methyl group.

TABLE IV

STOICHIOMETRY OF THYMIDYLATE FORMATION FROM DEOXYURIDYLATE[a,b]

| Incubation period | Increase of absorbancy at 340 mμ | Dihydrofolate formed (mμmole) | Thymidylate formed (mμmole) | Dihydrofolate / Thymidylate |
|---|---|---|---|---|
| 10 | 0.100 | 17.2 | 19.2 | 0.89 |
| 20 | 0.180 | 28.2 | 28.2 | 1.00 |
| 30 | 0.211 | 33.0 | 33.4 | 0.99 |
| 40 | 0.225 | 35.3 | 36.0 | 0.98 |
| 50 | 0.232 | 36.3 | 37.5 | 0.97 |
| 60 | 0.235 | 36.7 | 40.0 | 0.92 |

[a] From Wahba and Friedkin (1961).

[b] The incubation mixture consisted of 0.040 μmole of dUMP (labeled with $P^{32}$, 80,000 cpm); 0.15 μmole of *dl*, L-tetrahydrofolate; and enzyme (0.2 mg of protein) in 1 ml of a mixture containing tris-HCl buffer, pH 7.4, 40 μmoles; 2-mercaptoethanol, 100 μmoles; ethylenediaminetetraacetate, 1.0 μmole; $MgCl_2$, 25 μmoles; and formaldehyde, 15 μmoles.

In addition to its role as carrier of the 1-carbon unit, tetrahydrofolic acid has been implicated as the reducing agent in the methylation reaction (Friedkin, 1959a; Humphreys and Greenberg, 1958). Direct evidence for this has been presented by Friedkin (1959a), who has reported the transfer of tritium from tritiated tetrahydrofolic to thymidylic acid. In addition, Wahba and Friedkin (1961) have recently described the stoichiometric oxidation of tetrahydrofolic to dihydrofolic acid during the synthesis of thymidylate from deoxyuridylate. For each mole of thymidylate formed, one mole of dihydrofolate appeared. These results are reported in Table IV. Friedkin (1959b) has proposed a scheme for the methylation of deoxyuridylate by $N^5$–$N^{10}$ methylenetetrahydrofolic acid. In Fig. 2, deoxyuridylate is seen to cleave the methylene group at $N^5$ with the formation of a labile intermediate between this methylene group and the 5-position of the uracil ring. An intramolecular oxidation-reduction would yield thymidylate and dihydrofolate in equal amounts.

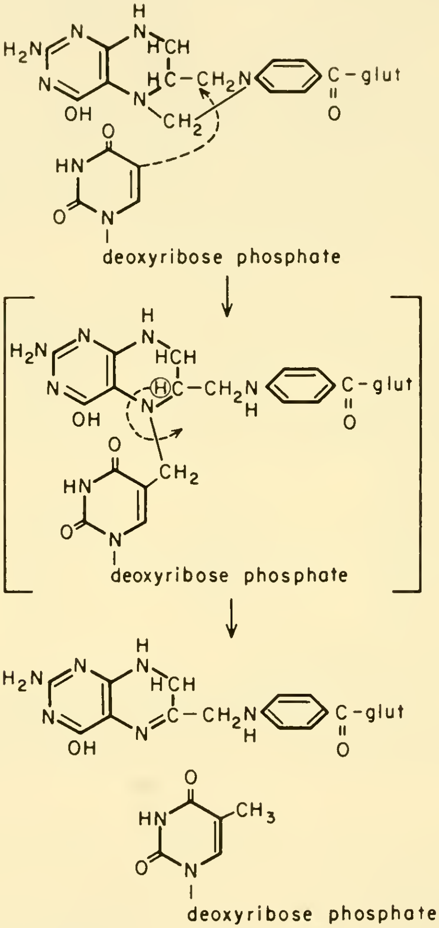

H₂N
OH
C–glut
deoxyribose phosphate
deoxyribose phosphate
deoxyribose phosphate

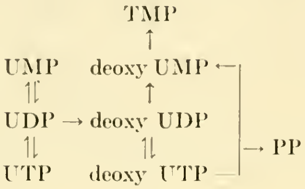
FIG. 2. Postulated mechanism for the synthesis of thymidylate. (From Friedkin, 1959b.)

Humphreys and Greenberg (1958) independently proposed a similar mechanism for this reaction.

An interesting aspect of the synthesis of thymidylic acid is the increase in the capacity of *E. coli* infected with certain bacteriophages to convert deoxyuridylate to thymidylate. Flaks and Cohen (1957) originally reported an increase in this activity in extracts of cells infected with bacteriophage T6 and suggested the name "thymidylate synthetase" since it was not known whether one or more enzymes was required to transfer the 1-carbon unit and carry out the reductive step (Flaks and Cohen, 1959b). These same investigators (1959b) demonstrated a 7-fold increase in thymidylate synthetase after infection with bacteriophage T2 and showed that the conversion of deoxyuridylate to thymidylate required formaldehyde and tetrahydrofolic acid; it did not proceed through 5-hydroxymethyldeoxyuridylate as an intermediate. Earlier experiments of Barner and Cohen (1954), in which they showed that a thymine-requiring strain of *E. coli*, $15T^-$, could synthesize thymine after infection with bacteriophage, can be explained by later experiments of these same authors (1959) in which they demonstrated that thymidylate synthetase (missing in $15T^-$) is induced by bacteriophage infection. This would support the idea that the infecting phage particle carries in the information for the induction of thymidylate synthetase, and Cohen (1961) investigated whether the enzyme formed after infection is the same as the enzyme in the normal cell. A comparison of thymidylate synthetase in extracts of *E. coli* B, $B_{T^-}$ infected with T2 and $B_{T^-}$ infected with T5 showed that the enzymes from all three sources were inhibited to the same extent by varying concentrations of 5-fluorodeoxyuridylate or D-arabinofuranosyl-5-fluorouracil-5′-phosphate, two powerful inhibitors of thymidylate synthetase. The active sites of the enzymes from these three sources are thus indistinguishable by this test. Recently, however, Greenberg *et al.* (1962b) have demonstrated that thymidylate synthetase from T2-infected *E. coli* has different chromatographic properties from the thymidylate synthetase of uninfected cells.

Cohen *et al.* (1958) have shown that 5-fluorodeoxyuridylate is a powerful non-competitive inhibitor of thymidylate synthetase, and at a concentration of $2 \times 10^{-6}\,M$ (a ratio of substrate to inhibitor of 1000) from 35–55% inhibition was observed. Heidelberger *et al.* (1957) have reported the anti-tumor activity of 5-fluorouracil and presented evidence for the inhibition of thymine synthesis in whole cells by this analog. The inhibition of thymidylate synthetase by 5-fluorodeoxyuridylate could explain this observation.

As mentioned previously (Section I,A), Bertani *et al.* (1961) have indicated that the conversion of uridylate to deoxyuridylate occurs at the diphosphate level. The bulk of the evidence indicates that the methylation of deoxyuridylate to thymidylate occurs at the monophosphate level. This would mean that deoxyuridine diphosphate would have to be converted to deoxyuridine monophosphate before methylation. Bertani *et al.* (1961) have found that deoxy-UDP is phosphorylated to deoxy-UTP in their system and that a specific UTPase is present which converts deoxy-UTP to deoxy-UMP and pyrophosphate. They suggest the following scheme for synthesis of deoxythymidylate in *E. coli:*



Another reaction which can provide deoxy-UMP directly for deoxy-TMP synthesis is the deamination of deoxycytidylic acid. Scarano (1958) has described a specific deoxycytidylate deaminase in sea urchin eggs and rabbit (Scarano and Tolarico, 1959) and Maley and Maley (1959) have extended this work to rat tissues and chick embryo. Keck *et al.* (1960) and Flaks and Cohen (1959b) reported the induction of a specific deoxycytidylate deaminase in T2-infected *E. coli.* Thus, in any of these systems the following pathway is possible:

$$\mathrm{CDP} \rightarrow \mathrm{dCDP} \rightleftharpoons \mathrm{dCMP} \xrightarrow{\text{deaminase}} \mathrm{dUMP} \rightarrow \mathrm{TMP}$$

Friedkin (1959b) has published an interesting account of the unique position of thymidylic acid in nucleotide metabolism.

Recently, Mandel and Borek (1961) reported that the methyl group of the thymine isolated from the ribonucleic acid of *E. coli* arises from the methyl group of methionine. This interesting observation would suggest that the small amount of thymine in RNA and the thymine in DNA are synthesized by two discrete pathways.

### C. SYNTHESIS OF 5-HYDROXYMETHYLDEOXYCYTIDYLIC ACID (dHMP)

The discovery that cytosine is completely replaced by the unique pyrimidine, 5-hydroxymethylcytosine (HMC), in the DNA of the T-even bacteriophages (Wyatt and Cohen, 1953) prompted an investigation of the enzymatic steps leading to the synthesis of this compound. Tracer experiments with whole cells had implicated cytosine as a precursor of the pyrimidine ring and the β-carbon of serine as a precursor of the hydroxymethyl group (Cohen and Weed, 1954) and eliminated methionine as the one-carbon donor (Green and Cohen, 1957). Flaks and Cohen (1957, 1959) first reported the synthesis of hydroxymethyldeoxycytidylate in cell-free extracts of $T6r^+$-infected *E. coli* and showed that analogous to thymidylate synthetase the system required tetrahydrofolic acid and formaldehyde. Tetrahydrofolic acid appears to act catalytically since the amount of dHMP formed is greater than the amount of tetrahydrofolate present. Dihydrofolic acid is inactive. The folic acid derivative presumably acts only as a one-carbon carrier in this system and has no function as a reducing agent as it does in thymidylate synthesis. Friedkin (1959b) has presented a hypothetical scheme for the reaction between deoxycytidylate and formyl tetrahydrofolic acid (Fig. 3). This scheme differs from the mechanism proposed for thymidylate (Fig. 2) in that no intramolecular shift of hydrogens occurs and hydrolysis of the intermediate leads directly to the hydroxymethyl derivative and free tetrahydrofolate. Deoxycytidylate hydroxymethylase, as

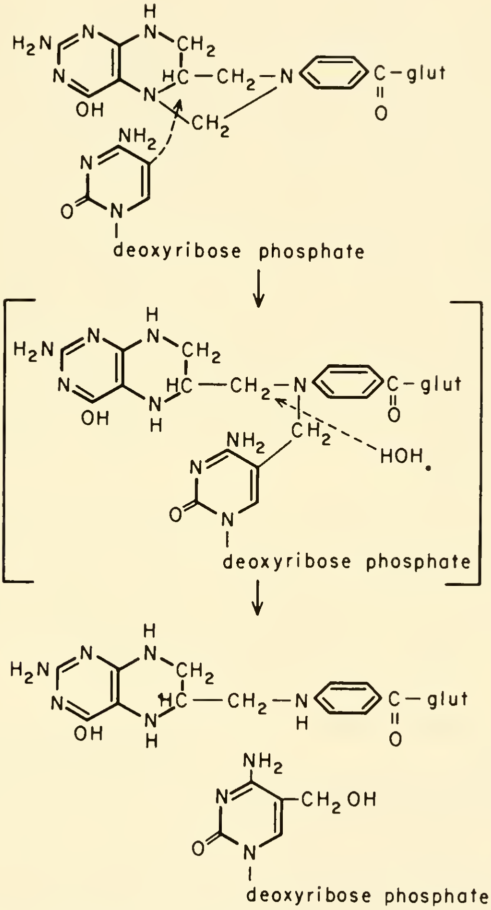


FIG. 3. Postulated mechanism for the synthesis of 5-hydroxymethyldeoxycytidylate. (From Friedkin, 1959b.)

the enzyme is called, is specific for dCMP and will not react with CMP, deoxycytidine, cytidine, or cytosine. It is distinct from thymidylate synthetase since it is much more stable in crude extracts of infected cells than the latter enzyme and it is not inhibited by 5-fluorodeoxyuridylate.

In respect to its distribution, deoxycytidylate hydroxymethylase has only been found in significant levels in extracts of *E. coli* infected with T2, T4, or T6 bacteriophages. It is not detectable in uninfected cells, in the viruses, or in cells infected with a virus which does not contain 5-hydroxymethylcytosine. Recently, Pizer and Cohen (quoted in Cohen, 1961), using very highly labeled formaldehyde in the enzyme assay, have shown that less than 1 molecule of active deoxycytidylate hydroxymethylase is present per uninfected bacterial cell. These calculations are based on a molecular weight of 64,000 ± 9000, and a turnover number of 1200 for the enzyme. It would appear then, that deoxycytidylate hydroxymethylase is either present in uninfected bacterial cells in an inactive form and is activated after infection or its formation is induced by injection of the viral DNA. Experiments designed to test these two alternatives tend to favor the latter, although no definitive statement can be made.

## III. PHOSPHORYLATION OF DEOXYRIBONUCLEOTIDES

### A. DEOXYRIBONUCLEOSIDE MONOPHOSPHATE KINASES

The name deoxyribonucleotide kinase has been given to a group of enzymes which catalyze the following reaction:

$$\text{Deoxyribonucleoside monophosphate} + \text{ATP} \rightleftharpoons \text{deoxyribonucleoside diphosphate} + \text{ADP}$$

This classification is misleading because it ascribes a degree of specificity to this group of enzymes which is not warranted. Both of the enzymes in this category, which have been purified and studied in detail (deoxycytidylate and deoxyguanylate kinases), catalyze the phosphorylation of ribo- as well as deoxyribonucleotides.

The occurrence of phosphorylated deoxyribonucleotides in animal tissues was reported by Potter *et al.* (1957), who identified the di- and triphosphates of deoxythymidine and deoxycytidine in the acid-soluble fraction of thymus glands. Interestingly enough, these authors found no evidence for the polyphosphates of the corresponding purine deoxyribonucleosides. Sable *et al.* (1954) observed that rat liver mitochondria actively oxidizing glutamate could form dADP and dATP from dAMP. The demonstration of adenylate kinase in liver mitochondria (Siekevitz and Potter, 1953) suggests that Sable's observation is probably related to the reports of Klenow and Anderson (1957) and Klenow and Lichtler (1957) that adenylate kinase (myokinase) can use adenine deoxyribonucleotides in place of ribonucleotides. This possibility is supported by the observation of Lehman *et al.* (1958) that heated extracts of *E. coli* partially retain their ability to phosphorylate deoxyadenylate while losing their capacity to phosphorylate other deoxyribonucleotides. Heat stability is a characteristic property of muscle adenylate kinase (Colowick and Kalckar, 1943). Hecht *et al.* (1954) reported the phosphorylation of dCMP to dCDP and dCTP in a cytoplasmic fraction of rat liver and found that dTTP was formed from dTMP in extracts of regenerating liver. Ochoa and Heppel (1957) have observed the phosphorylation of dAMP and dCMP in extracts of *Azotobacter*. Bollum (1958) and Canellakis and Mantsavinos (1958) reported the enzymatic synthesis of all four deoxyribonucleoside triphosphates in extracts of regenerating rat liver, and Kielley (1961) has studied these reactions in mouse liver.

It is interesting that normal rat liver, which can catalyze the phosphorylation of dAMP, dCMP, and dGMP by ATP, has at most a very low capacity to phosphorylate dTMP. The rapid increase in thymidylate kinase activity in regenerating rat liver has been demonstrated in several laboratories (Bollum and Potter, 1959; Canellakis, 1959; Canellakis *et al.*, 1959; Smellie, 1960). This increase is inhibited by post-operative X-irradiation of the whole animal (Bollum *et al.*, 1960). Weissman *et al.* (1960) and Smellie (1960) have studied the sequence of reactions:

$$\text{Deoxythymidine} \rightarrow \text{dTMP} \rightarrow \text{dTDP} \rightarrow \text{dTTP}$$

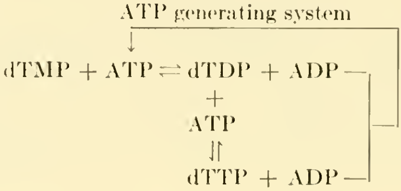
They claim that there is a sequential induction or appearance of thymidine kinase, thymidylate kinase, and thymidine diphosphate kinase in regenerating rat liver and in an L-strain of fibroblasts grown in tissue culture.

Gray *et al.* (1960) and Smellie (1960) claim there is a factor in normal rat liver which inhibits thymidylate kinase when added to extracts of regenerating liver. It is heat-labile, nondialyzable, and has no effect on dAMP, dGMP, and dCMP kinases. It will be interesting to learn more of this inhibitor which could play a role in regulating dTTP formation.

Bianchi *et al.* (1961) have reported that dTDP is not an intermediate in dTTP synthesis in mouse lymph nodes and leukemic spleen. They suggest that a pyrophosphate group is transferred from ATP according to the following reaction:

$$\text{dTMP} + \text{ATP} \rightarrow \text{dTTP} + \text{AMP}$$

This would be the only instance in the synthesis of deoxyribonucleoside triphosphates in which a transfer of the intact pyrophosphate supposedly occurs, and the reaction should be studied with the purified enzyme. The small amount of dTDP found in the chromatograms (Bianchi *et al.*, 1961) was attributed to the hydrolysis of dTTP. The other possibility (which I favor) is that dTDP is an intermediate but does not accumulate in the reaction mixture because another enzyme, nucleoside diphosphate kinase, is present in the crude extract which converts dTDP to the triphosphate. This would be in keeping with their conditions of incubation which include an ATP-generating system. The continued conversion of ADP to ATP by the generating system would tend to drive the reaction to the right, hence favoring the formation of dTTP and decreasing the accumulation of dTDP:



We have evidence (M. Rienhoff and M. Bessman, in preparation) that partially purified extracts of regenerating rat liver catalyze the phosphorylation of dTMP to the corresponding diphosphate.

The interesting observation (Hiatt and Bojarski, 1960) that intraperitoneal injections of thymidine stimulate the formation of thymidylate kinase in rat liver was explained by these same authors (Bojarski and Hiatt, 1960) as a stabilization of thymidylate kinase during preparation of the extract. They showed that preincubation of a liver extract resulted in a rapid decrease of thymidylate kinase activity which could be prevented by including thymidylate (or thymidine at 10 times the concentration) in the preincubation medium. This information has been useful in preparing extracts for further fractionation.

Kornberg (1957) and Lehman *et al.* (1958a) described the phosphorylation of dAMP, dCMP, dGMP, and dTMP by a partially purified extract from *E. coli*. It became apparent that more than one enzyme was responsible for the phosphorylation of these four different nucleotides when Lehman *et al.* (1958a) observed that heated extracts lost their ability to phosphorylate dCMP, dGMP, and dTMP while partially retaining their ability to phosphorylate dAMP. An indication that the other three activities were due to separate enzymes was the observation of Hurwitz (1959) that acetone fractionation resulted in relative enrichments of dCMP, dGMP, and dTMP kinases in different fractions.

The only report of the study of a purified kinase to date is that of

Maley and Ochoa (1958), who described the properties of the deoxycytidylate kinase in *Azotobacter vinelandii.* This enzyme was purified 800-fold over the crude extract and was shown to catalyze the following reaction:

$$\text{dCMP} + \text{ATP} \rightleftharpoons \text{dCDP} + \text{ADP}$$

The other three nucleotides (dGMP, dTMP, or dAMP) are not substrates for this enzyme.

The stoichiometry of the reaction is shown in Table V. In Expt. 1, a less purified enzyme was used and it can be seen that both dCDP and dCTP were formed probably due to the presence of a nucleoside diphosphate kinase which is known to be present in *Azotobacter.* In the more purified fractions (Expts. 2 and 3) only the diphosphate was formed. The same enzyme appears to catalyze the phosphorylation of the ribo compound, cytidylic acid, since the ratio of the activities toward the two substrates remains constant during the purification. Cytidylate is actually phosphorylated at about 1.5 times the rate of deoxycytidylate by this enzyme at equal concentrations of the two compounds.

TABLE V

STOICHIOMETRY OF REACTION: ATP + dCMP ⇌ ADP + dCDP[a,b]

| Expt. No. | ATP | dCMP | ADP | dCDP | dCTP |
|---|---|---|---|---|---|
| 1 | −4.72 | −4.35 | +4.28 | +3.05 | +1.15 |
| 2 | −4.54 | −3.79 | +4.24 | +3.08 | 0 |
| 3 | +0.71 | +0.70 | −0.70 | −0.73 | 0 |

[a] From Maley and Ochoa (1958).

[b] All samples contained 100 μmoles of tris buffer, pH 7.5, and 20 μmoles of $MgCl_2$ in a final volume of 1.0 ml and were incubated at 30°C. The remaining components and conditions were as follows: Expt. 1: 9 μmoles of ATP, 9 μmoles of dCMP, and 10 μg of enzyme of specific activity 60 with incubation for 15 hours; Expt. 2: 10 μmoles of ATP, 5.8 μmoles of dCMP, and 2.3 μg of enzyme of specific activity 175 with incubation for 3 hours; Expt. 3: 4.5 μmoles of ADP, 1.1 μmoles of dCDP, and 2.3 μg of enzyme of specific activity 175 with incubation for 18 hours. Results are expressed as μmoles decrease or increase in reactants and products.

We have purified a deoxyguanylate kinase (M. Bessman, unpublished) from extracts of *E. coli* approximately 400-fold which appears to catalyze the following reaction.

$$\text{dGMP} + \text{ATP} \rightleftharpoons \text{dGDP} + \text{ADP}$$

Ion exchange chromatography reveals that more than 95% of the phosphorylated product is the diphosphate. Again, ribo-GMP is also phosphorylated but dCMP and dTMP are not.

Both these purified enzymes (deoxycytidylate and deoxyguanylate kinases) phosphorylate the ribo- as well as deoxyribonucleotide. The specificity seems to reside in the structure of the base and does not involve the sugar moiety. The name deoxynucleotide kinase is thus not appropriate in its strictest sense.

An interesting feature of the deoxynucleotide kinases of *E. coli* is that they are markedly influenced by bacteriophage infection. Activities of dTMP and dGMP kinases rapidly increase after T2, T4, or T6 infection, and dTMP, dGMP, and dCMP kinases increase after T5 infection (Kornberg *et al.*, 1959; Bessman, 1959). In addition, a new kinase, undetectable in the uninfected cell, which catalyzes the phosphorylation of hydroxymethyldeoxycytidylate is formed after infection with T2, T4, and T6 bacteriophages (Kornberg *et al.*, 1959; Somerville *et al.*, 1959). Recently it has been shown that the dGMP and dTMP kinases formed after bacteriophage infection are different from their counterparts in the normal cell (Bessman and Van Bibber, 1959; Bello *et al.*, 1961a,b). They can be separated from the kinases of the uninfected cell by chromatography on DEAE cellulose columns and have different properties. For example, dGMP kinase from T2-infected *E. coli* is absolutely specific for the deoxyribonucleotide. It will not phosphorylate ribo-GMP, whereas the enzyme from the uninfected cell will.

Purification of dGMP kinase from T2-infected cells has led to a concomitant purification of dTMP and 5-hydroxymethyldeoxycytidylate kinases (Bessman and Bello, 1961). Since the ratios of the three activities do not change over a 200-fold range in purification these authors investigated the possibility that all three activities resided in the same protein. Results of heat-inactivation studies and competitive inhibition measurements have led to the tentative conclusion that the three kinases are associated with one protein, although the data can be interpreted equally well by assuming the activities are associated with three very closely related proteins. Preliminary studies with the T5-induced kinases show that the dCMP, dGMP, and dTMP kinases, which increase after infection, chromatograph in the same place on columns of DEAE cellulose, suggesting a close relationship if not identity of these three activities. Again, the phage-induced dCMP kinase differs from the dCMP kinase of normal cells by requiring deoxyribonucleotides as substrates. Ribo-GMP or CMP are not phosphorylated. Thus, in this case and in the T2-infected cell, the name "deoxyribonucleotide" kinase is truly applicable. These are the only two members of the so-called deoxyribonucleotide kinases which are specific for the deoxy compounds.

### B. DEOXYNUCLEOSIDE DIPHOSPHATE KINASES

The phosphorylation of deoxyribonucleotides most likely proceeds through the corresponding diphosphates to the triphosphate level. In the two cases cited above which were studied with purified enzymes (dCMP and dGMP kinases from *E. coli*) the diphosphate accumulates in the reaction mixture and no triphosphate is formed. However, as discussed above, crude extracts of *E. coli*, regenerating rat liver, leukemic cells, Ehrlich ascites cells, thymus tissue, etc., form predominantly triphosphates in the presence of ATP and a deoxyribonucleoside monophosphate. These two observations taken together mean there must be an enzyme(s) which catalyzes the phosphorylation of the deoxyribonucleoside diphosphates formed as intermediates during the incubation. However, no specific deoxyribonucleoside diphosphate kinases have been reported, with the possible exception of the dTDP kinase alluded to in the report of Weissman *et al.* (1960). It is not unlikely that the nucleoside diphosphate kinases described by Berg and Joklik (1953) and Krebs and Hems (1953), which catalyze the phosphorylation of the ribonucleoside diphosphates, also react with the corresponding deoxy compounds:

$$\text{(deoxy)nucleoside diphosphate} + \text{ATP} \rightleftharpoons \text{(deoxy)nucleoside triphosphate} + \text{ADP}$$

Kirkland and Turner (1959) have shown that, in addition to the phosphorylation of IDP and UDP originally demonstrated by the above investigators, purified extracts of peas also phosphorylate CDP and GDP.

Davidson (1959), in a reinvestigation of the specificity of pyruvate kinase, has demonstrated that GDP can replace ADP in the reaction. It is possible that dGDP also could be phosphorylated in this system since Klenow and Anderson (1957) presented evidence for the replacement of ADP by dADP using crystalline pyruvate kinase from rabbit muscle. It is clear that the specificity of these enzymes with respect to the deoxy compounds should be investigated.

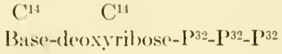
## IV. SYNTHESIS OF DNA BY *Escherichia coli* POLYMERASE

### A. MEASUREMENT OF DNA SYNTHESIS

#### 1. *Acid Insolubility*

The insolubility of DNA in dilute acids affords a convenient means of separating the macromolecule from low molecular weight deoxynucleotide derivatives which are soluble in acid solution. For the ensuing discussion, we shall define "DNA synthesis" to mean the incorporation of acid-soluble deoxyribonucleotide derivatives into covalent linkage in an acid-insoluble product. This is purposely a broad definition which includes "exchange reactions" if they occur as well as reactions in which there is a net synthesis of polymer. Arguments for a narrower definition have an emotional rather than practical basis.

The most sensitive measure of DNA synthesis involves deoxyribonucleotides labeled with radioactive isotopes of carbon or phosphorus.



$$\overset{C^{14}}{\text{Base}}\text{-}\overset{C^{14}}{\text{deoxyribose}}\text{-}P^{32}\text{-}P^{32}\text{-}P^{32}$$

Any combinations of these labels may be present in the substrate, but it should be borne in mind that the two terminal phosphorus atoms will not become part of the DNA molecule. This technique is extremely sensitive and has been used to measure very minute reactions. For example, the incorporation of 1 $\mu\mu$mole of deoxyribonucleotide with a specific activity of $1 \times 10^8$ cpm per $\mu$mole would give 100 cpm above background. By suitable wash procedures, the blank values can be kept to a relatively low level. For routine assays, specific activities two orders of magnitude lower ($1 \times 10^6$ cpm per $\mu$mole) may be employed.

For situations in which substantial amounts of DNA have been synthesized, the acid-insoluble precipitate may be used for direct measurements of deoxyribose, phosphorus, or ultraviolet extinction.

### 2. *Viscosity Measurements*

"The viscosity of dilute DNA solutions exceeds that of any other widely studied natural or synthetic substance. . . ." (Doty, 1957). Pouyet

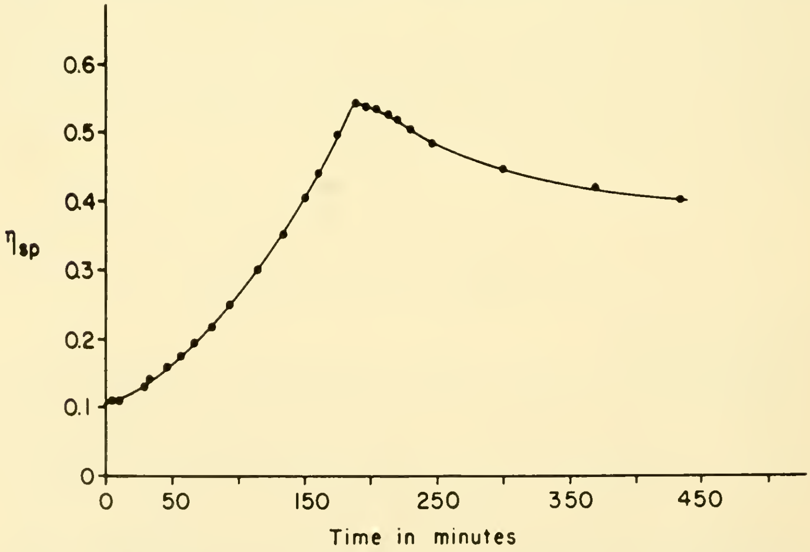


FIG. 4. Measurement of DNA synthesis by viscosity. (From Lehman, 1959.)

(1952) obtained a value of 50 deciliters (dl)/gm for the intrinsic viscosity using a Couette viscometer, and several other measurements of various DNA samples put the range between 40 to 60. This property is very useful for kinetic measurements of DNA synthesis since the whole reaction may be carried out in a viscometer. Figure 4 demonstrates the increase in specific viscosity of a reaction mixture primed with calf thymus DNA. This technique is applicable to systems in which the newly synthesized DNA represents a substantial increase over the initial DNA content of the reaction mixture.

### 3. *Measurements of Hypochromicity*

Kunitz (1950) observed an increase in the ultraviolet extinction of solutions of DNA during enzymatic digestion with deoxyribonuclease and utilized this "hyperchromic effect" as a means to assay this enzyme. This hyperchromic effect is due to the fact that the ultraviolet extinction of a nucleic acid preparation is less than the sum of the extinctions of its individual nucleotides. The theoretical basis for this effect is under in-

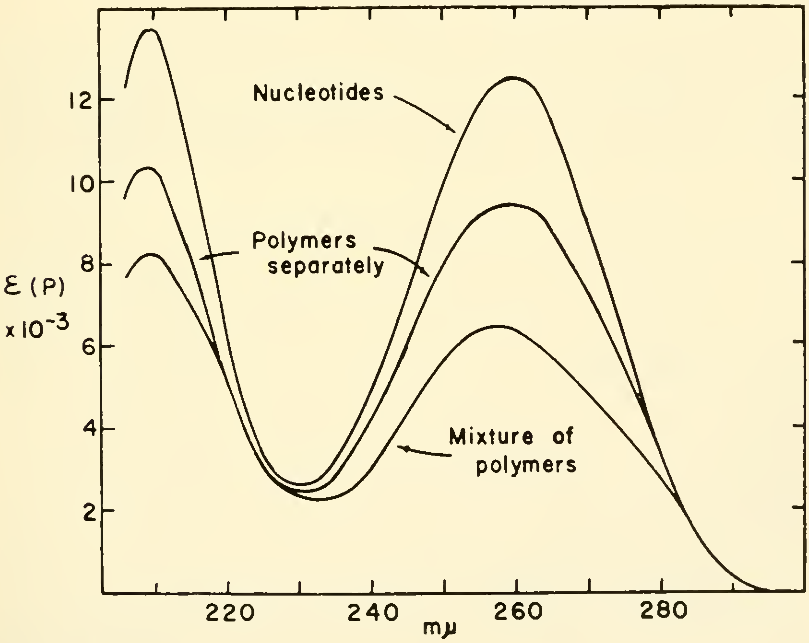


FIG. 5. Absorption spectra of an equimolar mixture (phosphorus basis) of polyadenylic acid and polyuridylic acid. The upper curve refers to the mononucleotides obtained by alkaline hydrolysis of the mixture. The middle curve is that for the separately measured polymer spectra. The lower curve is the measured curve for the mixture of polymers. (From Warner, 1957.)

vestigation (cf. Shugar, 1960) but it is clear that there are two distinct contributions to the over-all hypochromicity of DNA solutions. The first is maintained through the hydrogen bonding between the base pairs in the twin-stranded Watson-Crick model and accounts for approximately 50% of the total hypochromicity. The remainder is maintained through the phosphodiester bridges between the individual nucleotides (Sinsheimer, 1954; de Garilhe and Laskowski, 1956; Michelson, 1958). An illustration of this is shown in Fig. 5.

The decrease in extinction during polymerization of mononucleotides can be made the basis of an assay for DNA synthesis. The method is not suitable for measurements of small increments of DNA since the effect is relatively small and in addition to the absorption of the nucleotides themselves there is the contribution of the primer DNA. The method has been used successfully in measuring unprimed reactions and reactions in which extensive synthesis of polymer occurs (see Section IV,G).

### B. REQUIREMENTS OF THE REACTION

#### 1. *Forward Reaction*

The synthesis of DNA catalyzed by purified fractions of *E. coli* polymerase requires the presence of all four of the deoxyribonucleoside triphosphates commonly found in DNA, Mg ions, and a highly polymerized preparation of DNA (Bessman *et al.*, 1958a). In Table VI, the effect of omitting any one or combinations of these constituents is reported. Within the limits of this particular assay, omission of any one

TABLE VI

REQUIREMENTS FOR DEOXYNUCLEOTIDE INCORPORATION INTO DNA[a,b]

| System | m$\mu$moles |
|---|---|
| Complete system | 0.50 |
| Omit dCTP, dGTP, dATP | <0.01 |
| Omit dCTP | <0.01 |
| Omit dGTP | <0.01 |
| Omit dATP | <0.01 |
| Omit $Mg^{++}$ | <0.01 |
| Omit DNA | <0.01 |
| DNA pretreated with DNase | <0.01 |

[a] From Bessman *et al.* (1958a).

[b] The complete system contained 5 m$\mu$moles of dTTP (dTP$^{32}$PP, 1.5 × $10^6$ cpm per $\mu$mole), dATP, dCTP, and dGTP, 1 $\mu$mole of $MgCl_2$, 20 $\mu$moles of glycine buffer, pH 9.2, 10 $\mu$g of DNA, and 3 $\mu$g of "polymerase" fraction V in a final volume of 0.30 ml. The incubation was carried out at 37°C for 30 minutes.

of the deoxyribonucleoside triphosphates of adenine, cytosine, or guanine prevents DNA synthesis. Also, there is an absolute requirement for Mg ions and DNA. Preincubation of the DNA with DNase (5 $\mu$g, 37°C, 30 minutes) completely destroyed its priming ability. In a similar experiment, using a different $P^{32}$-labeled nucleotide as an index of DNA synthesis, it could be shown that dTTP is also required.

### 2. *Reverse Reaction*

Concentrations of inorganic pyrophosphate equal to that of the individual triphosphates ($1.6 \times 10^{-5}\ M$) have no effect on the synthetic reaction rate. However, at pyrophosphate concentrations 200-fold higher ($3 \times 10^{-3}\ M$) the reaction rate is inhibited 50% and pyrophosphate is incorporated into deoxyribonucleoside triphosphates, as shown in Table VII. Unlike the synthetic reaction the reversal proceeds well in the absence of three of the four triphosphates. Omission of all four, however, reduces the rate of pyrophosphorolysis considerably.

TABLE VII

EVIDENCE FOR REVERSAL OF REACTION[a]

| System | $P^{32}$-deoxynucleoside triphosphate[b] (m$\mu$moles) |
|---|---|
| Complete system | 3.47 |
| Omit DNA | <0.02 |
| Omit enzyme | <0.02 |
| Omit dATP | 2.34 |
| Omit dATP, dTTP | 2.14 |
| Omit dATP, dTTP, dGTP | 1.34 |
| Omit dATP, dTTP, dGTP, dCTP | 0.20 |
| Replace $P^{32}P^{32}$ with $Pi^{32}$ | <0.02 |
| Replace DNA with DNase-treated DNA | <0.02 |

[a] From Bessman *et al.* (1958a).

[b] These values are corrected for a zero-time value of 0.05 m$\mu$mole. The complete system contained 5 m$\mu$moles each of dATP, dTTP, dGTP, and dCTP, 1 $\mu$mole of $P^{32}P^{32}$ ($5 \times 10^5$ cpm $\mu$mole), 10 $\mu$g of DNA, 2 $\mu$mole of $MgCl_2$, 20 $\mu$moles of tris buffer, pH 7.5, and 0.8 $\mu$g of "polymerase" fraction VII in a final volume of 0.3 ml. After incubation at 37°C for 30 minutes the mixture was chilled and cold solutions of albumin (0.05 ml of 10 mg per ml), perchloric acid (0.25 ml of 1 $N$), and "carrier" PP (0.5 ml of 0.1 $M$, pH 7) were added. The nucleotides were adsorbed on charcoal which was then washed, plated, and counted.

That pyrophosphate is actually a product of the reaction is shown in Table VIII. For each mole of dTMP incorporated into DNA, one mole of pyrophosphate is released. The stoichiometry of this reaction and the

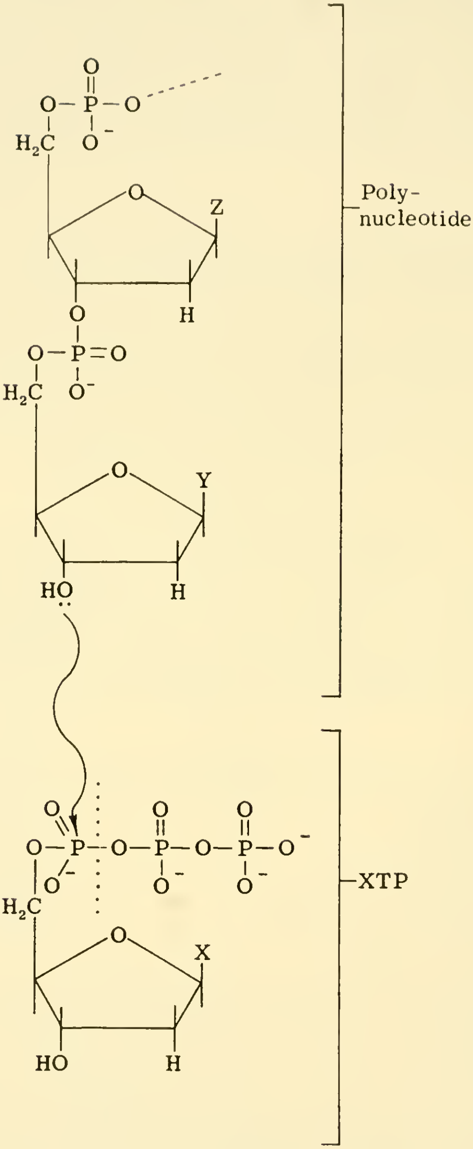

Poly-
nucleotide
XTP
Z
Y
X

FIG. 6. Postulated mechanism for extending a DNA chain. (From Kornberg, 1960.)

TABLE VIII

LIBERATION OF INORGANIC PYROPHOSPHATE[a,b]

| Estimation | Control[c] (mμmoles) | Experimental (mμmoles) | Δ (mμmoles) |
|---|---|---|---|
| DNA-C$^{14}$ incorporation | 1 | 25 | 24 |
| P$^{32}$P$^{32}$ released | 2 | 22 | 20 |
| Pi$^{32}$ released | 7 | 11 | 4 |

[a] From Bessman *et al.* (1958a).

[b] The labeled substrate was dTP$^{32}$P$^{32}$P-2-C$^{14}$.

[c] dATP was omitted.

requirements for the system have led to the following equation for the synthesis of DNA (Kornberg, 1960):

$$\begin{matrix} n\,\text{TPPP} \\ n\,\text{dGPPP} \\ n\,\text{dAPPP} \\ n\,\text{dCPPP} \end{matrix} + \text{DNA} \rightleftharpoons \text{DNA} - \begin{bmatrix} \text{dTP} \\ \text{dGP} \\ \text{dAP} \\ \text{dCP} \end{bmatrix}_n + 4\,(n)\,\text{PP}$$

The mechanism of the reaction is viewed by Kornberg (1960) as the nucleophilic attack of the terminal 3′-hydroxyl group of the DNA chain on the nucleotidyl phosphorus atom of the deoxyribonucleoside triphosphate with the elimination of pyrophosphate. By this mechanism, the DNA chain is lengthened from the nucleoside and as depicted in Fig. 6. This mechanism is consistent with all the present data.

### C. THE LIMITED REACTION (Adler *et al.*, 1958)

In the previous section (IV,B,1), a requirement for all four deoxyribonucleoside triphosphates was shown to be necessary for DNA synthesis. However, using highly labeled substrates, a small but significant incorporation of single deoxyribonucleotides can be demonstrated (Table IX). This incorporation amounts to approximately 0.1% of the rate of the complete system and occurs with any of the four deoxyribonucleoside triphosphates. The reaction proceeds rapidly at first and then falls off, suggesting that there are preferred sites of attachment which become saturated. Several lines of evidence suggest that this reaction is not due to the presence of small amounts of the other deoxyribonucleotides in the enzyme preparation, in the added DNA, or in the labeled triphosphate; and the possibility that the other triphosphates are formed by pyrophosphorolysis of the DNA during the incubation has been ruled out.

When a product of this limited reaction (containing a labeled deoxyribonucleotide) is isolated, washed free of acid-soluble deoxyribonucleotides and hydrolyzed with snake venom phosphodiesterase, almost all

TABLE IX
INCORPORATION OF SINGLE DEOXYNUCLEOTIDES INTO DNA[a]

| Additions[b] | DNA-$P^{32}$ (mμmoles) |
|---|---|
| dCTP + dGTP + TTP + dATP | 3300 |
| dCTP + dGTP + TTP | 15.7 |
| dCTP + dGTP | 5.1 |
| dCTP | 2.5 |
| dCTP (DNA omitted) | 0.0 |

[a] From Kornberg (1959).

[b] The incubation mixture (0.3 ml) contained 5 mμmoles of $dCP^{32}PP$ ($7.1 \times 10^7$ cpm per μmole), and of the other deoxynucleoside triphosphates where indicated, 1 μmole of $MgCl_2$, 20 μmoles of glycine buffer, pH 9.2, 10 μg of DNA, and 3 μg of enzyme fraction VI. The incubation was carried out at 37°C for 30 minutes.

of the radioactivity is released before appreciable amounts of the DNA has been rendered acid-soluble (Fig. 7). Since this enzyme has been shown to hydrolyze polynucleotides (ribo- and deoxyribo-) in a stepwise fashion starting from the nucleoside end (Laskowski *et al.*, 1957; Bowman, 1957; Khorana *et al.*, 1958; Singer *et al.*, 1958), it would appear that the added deoxyribonucleotides are localized at the ends of the DNA chain. On the average, one deoxyribonucleotide is added per DNA chain. It may be asked whether each of the deoxyribonucleotides is added to specific DNA chains terminating with a specific base or whether dCTP for example will add to DNA chains terminating in any of the deoxyribonucleotides. Results of nearest neighbor analyses

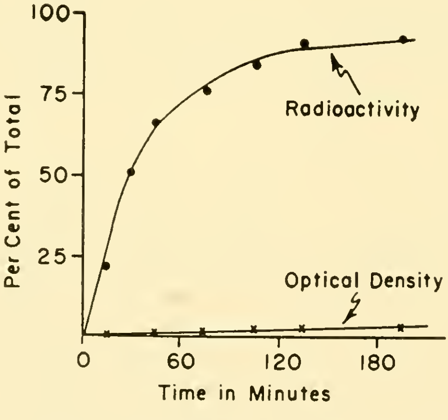


FIG. 7. Action of venom phosphodiesterase on limit product. (From Adler *et al.*, 1958.)

(see Section IV,F) favor the latter interpretation since it can be shown that each of the four deoxyribonucleotides attaches to chains ending in each of the bases. This analysis also rules out the possibility that the incorporation of single deoxyribonucleotides is due to the formation of homopolymers of the individual deoxyribonucleotides (poly-dCMP, for example), and demonstrates that the deoxyribonucleotide is attached to the DNA in covalent linkage through its phosphorus atom and not by some nonspecific adsorption or inclusion, etc.

What is the significance of the limited reaction? One possibility is that this is the prototype of the reaction that occurs when all four deoxyribonucleotides are present. If DNA replicates by adding deoxyribonucleotides to its end, these new units must loop back in some manner to allow the directing influence of the DNA chain to be exerted. But this imposes an additional requirement, that is some means of breaking this loop when the DNA has been duplicated. We know that in reactions in which there has been a 20-fold synthesis over the DNA originally present as primer, the molecular weight of the product is not increased 20-fold.

Another possibility is that the chains of the double helix of primer DNA as isolated from natural sources are of unequal length (possibly through damage during preparation) and the limited reaction is the repair of these broken ends. This could mean that the sequence of deoxyribonucleotides added to the shorter ends would be dictated by the longer end. In keeping with this notion is the recent observation (A. Kornberg, personal communication) that in the limited reaction using the d-AT polymer as primer (see Section IV,G) dATP or dTTP is incorporated, whereas dGTP and dCTP are not. This would indicate that there is some specificity imposed by the primer even for the limited additions to the ends of the chains. This "repair" mechanism would not be pertinent to a scheme for the replication of DNA which does not involve a covalent linkage at any stage between primer and product.

A third possibility is that the enzyme which carries out the limited reaction is different from *E. coli* polymerase. This possibility seems remote.

### D. NATURE OF THE PRIMER

So far, using highly purified fractions of *E. coli* polymerase, it has not been possible to synthesize a product containing all four deoxyribonucleotides in the absence of a primer. The priming DNA can be obtained from many sources, e.g., *E. coli,* calf thymus, salmon sperm, bacteriophage, crab testes, wheat germ, etc., but at present no chemically synthesized product has been able to substitute for the natural polymer.

As shown in Table VI, DNA treated with DNase loses its capacity to support synthesis. Other polymers which have been tested (Bessman *et al.,* 1957) and found ineffective as primers are apurinic acid, turnip yellow mosaic virus RNA, and a polythymidylic acid polymer (kindly provided by H. Gobind Khorana).

The important feature of the priming DNA seems to lie in its primary structure since many treatments designed to destroy the hydrogen bonding between the specific base pairs of double helices (Watson and Crick, 1953a) not only do not destroy priming ability but in some instances actually enhance it. In Table X, for example, native DNA is compared to heat-denatured DNA and "natural" single-stranded DNA (Sinsheimer, 1959). It can be seen that the "denatured" or single-stranded DNA primes synthesis at twice the rate and to the same extent as the natural DNA. Furthermore, the products of the reaction are in all cases viscous even though the heated, and single-stranded DNA are nonviscous preparations.

TABLE X

ENZYMATIC SYNTHESIS OF VISCOUS DNA FROM NON-VISCOUS PRIMERS[a]

| | Thymus DNA | Heated thymus DNA | $\phi$X174 DNA |
|---|---|---|---|
| Reaction | | | |
| Rate | 1 × | 2 × | 2 × |
| Extent | Tenfold | Tenfold | Tenfold |
| Viscosity (dl/gm) | | | |
| Primer | 45 | — | 1 |
| Product | 41 | 20 | 22 |
| Heated product | <1 | — | <1 |
| Hyperchromic effect (%) | | | |
| Primer | 30 | — | 10 |
| Product | 30 | — | 30 |

[a] From Lehman (1959).

An interesting observation (Lehman, 1959) is the improvement of calf thymus DNA as a primer by treating it with minute amounts of pancreatic DNase:

$$\frac{\mu\text{g DNA}}{\mu\text{g DNase}} = 2 \times 10^5$$

DNA treated in this manner primes syntheses at 2 to 3 times the rate of untreated primer. It is possible that this treatment provides more free 3′-hydroxyl groups for condensation with the deoxynucleoside triphosphates. Alternatively, the pretreatment could provide weak points throughout the molecule by hydrolysis of phosphodiester linkages, thus facilitating strand separation and enhancing priming efficiency.

With some highly purified fractions of *E. coli* polymerase, native DNA acts very poorly as a primer compared to heated DNA or DNA slightly hydrolyzed with DNase. This raised the question (Kornberg, 1960) of whether native DNA (double-stranded) can function at all as primer. To test this possibility rigorously, *E. coli* polymerase preparations free of nuclease activity must be used; unfortunately, at present, none are available.

## E. NATURE OF THE PRODUCT

### 1. *Chemical Structure*

The deoxynucleotides incorporated into DNA by *E. coli* polymerase are present in 3′–5′ phosphodiester linkage (Bessman *et al.*, 1958a). A product labeled with $C^{14}$-thymidylate was digested with deoxyribonuclease and chromatographed according to Sinsheimer (1954). Two dinucleotides, cytidylate-thymidylate and thymidyl-thymidylate, were isolated and characterized as dinucleotides with typical 3′–5′-phosphodiester bridges. Thus it is clear that covalent linkages are formed during the incorporation and the four deoxyribonucleotides add to each other. This rules out the possibility, for example, that homopolymers of the individual deoxyribonucleotides are the sole products. Besides having the normal internucleotide linkages, the product reacts with diphenylamine to give the characteristic color for deoxyribose (Dische, 1955). This is shown in Table XI, where the values for deoxyribonucleotide incorporation (as measured by $P^{32}$), ultraviolet extinction, and deoxyribose content of a synthetic DNA are compared.

TABLE XI
NET SYNTHESIS OF DNA[a]

| Expt. No. | Estimation | Control (no enzyme) | Complete | Δ (μmoles) |
|---|---|---|---|---|
| 1 | $P^{32}$ incorporation | 0.000 | 0.280 | 0.280 |
| | Optical density | 0.193 | 0.458 | 0.265 |
| | Deoxypentose | 0.187 | 0.395 | 0.208 |
| 2 | Optical density | 0.060 | 0.628 | 0.568 |
| 3 | Optical density | 0.052 | 0.583 | 0.531 |
| 4 | Optical density | 0.051 | 0.642 | 0.591 |
| 5 | Optical density | 0.041 | 0.889 | 0.848 |

[a] From Bessman *et al.* (1958a).

By these criteria, natural DNA and the product synthesized with *E. coli* polymerase are identical.

### 2. *Physical Structure*

Physical measurements of the DNA synthesized by *E. coli* polymerase became feasible when net synthesis of product was accomplished. In Table XI, several preparations of DNA are shown in which increases up to 20 times that of the added primer indicate that at least 95% of the total DNA represents newly synthesized material. These products were

TABLE XII

PHYSICAL PROPERTIES OF ENZYMATICALLY SYNTHESIZED DNA[a]

| | | | Heated at 100°C for 15 minutes | |
|---|---|---|---|---|
| Property | Primer | Product | Primer | Product |
| Sedimentation coefficient | 25 | 20–25 | 20 | 14 |
| Intrinsic viscosity (dl/gm) | 40–50 | 15–45 | <1 | <1 |
| Molecular weight | $8 \times 10^6$ | $4\text{–}6 \times 10^6$ | | |

[a] From Lehman (1959).

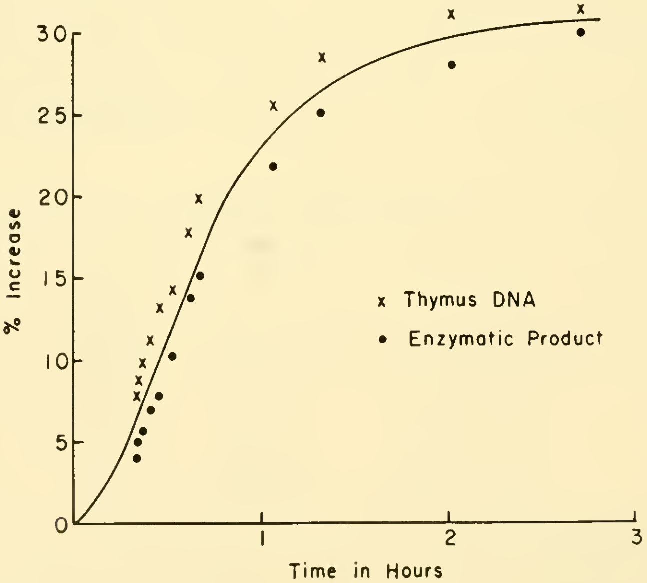


FIG. 8. Increase in ultraviolet absorption of DNA upon digestion with pancreatic deoxyribonuclease. (From Lehman, 1959.)

analyzed by Schachman *et al.* (1958) and in every parameter tested they resembled native calf thymus DNA. In Table XII the viscosities, sedimentation constants, and molecular weights for synthesized DNA and calf thymus DNA (used as primer) are compared, and it can be seen that the two materials are closely related. Molecular weights of several preparations range from 4 to 6 million and the viscosity and sedimentation data indicate that the deoxynucleotide units are organized in the macromolecule to form relatively stiff rods with effective volumes greater than would be expected from single polynucleotide chains with freedom of rotation at each internucleotide link. If the product is heated at 100°C for 15 minutes (Table XII) there is a sharp drop in viscosity with a relatively small decrease in sedimentation. When the product is digested with DNase there is a 30% increase in extinction, or hyperchromic effect (Fig. 8). These latter two observations support the concept of an organized rigid arrangement of deoxyribonucleotides in the superstructure of the native molecule.

### F. PRIMER-PRODUCT RELATIONSHIPS

As mentioned in the introduction, the term "replication" is taken to mean the synthesis of a DNA which is in every way indistinguishable from the primer. In this sense it would be better to use the word "template" rather than primer to describe the small amount of DNA required for the synthesis of more DNA since this implies a *copying* mechanism rather than *initiating* mechanism as the function of the added DNA although, the two functions go hand in hand, except in some instances mentioned later. "Primer" will be used throughout this discussion for convenience, but it should be borne in mind that the primer used in DNA synthesis is unique in biochemistry in that it serves both as initiator *and* template.

So far we have seen that in respect to gross chemical structure and physical profile the synthetic DNA closely resembles the primer. We would like to be able to determine the exact sequence of deoxyribonucleotides in the product and compare this to the primer molecule; but, at present, techniques are not available for doing this even for relatively small polynucleotides. Valuable information has been obtained toward elucidating the fine structure of the synthetic DNA by determining its base composition and dinucleotide frequencies.

#### 1. *Base Composition*

Two general methods have been used to determine the base composition of the newly formed DNA. In the first procedure four reaction mixtures identical except for the radioactive deoxyribonucleoside triphosphate are incubated separately, and the incorporation of label into the acid-insoluble product is measured in each. This method has the advantage of being applicable to very small syntheses and has been used to measure base ratios of products formed very early in the reaction. In the second method, products obtained from large-scale syntheses are hydrolyzed and chromatographed to determine base content directly. It is important that the primer represent only a small percentage of the total DNA analyzed so that routinely 10- to 20-fold net syntheses are employed in this procedure.

In Table XIII, the base compositions of several different DNA preparations are shown. First, in all cases, the percentage compositions of the bases in the products are very similar to these values for the respective primers. Second, the content of adenine approximates the content of thymine and the guanine and cytosine contents are about equal. Thus, purines equal pyrimidines. These relationships are in keeping with the proposed structure for DNA.

With the use of the radioactive assay it was possible to show that the composition of the product is established early in the reaction. In Table XIV it can be seen that when only 2% increase in DNA synthesis had occurred in the reaction primed with *Mycobacterium phlei* DNA, the ratio of thymine to cytosine incorporation was equal to that of the primer. This ratio remained constant with up to 163% increase in DNA. The same is true for the reactions primed with calf thymus and *Aerobacter aerogenes* DNA. The increase in ratios during the latter part of the reaction can be explained by the formation of a polymer of deoxyadenylate and deoxythymidylate formed late in the incubation. This polymer will be discussed in Section IV,G.

An important consideration in establishing what factor(s) is responsible in determining the ultimate composition of the product is whether the relative concentrations of the individual deoxyribonucleoside triphosphates in the incubation mixture influence their relative concentrations in the product. That this is not the case is indicated in Table XV, where it can be seen that although the relative concentration of dTTP and dATP were varied singly and together over a 5-fold range, the composition of the product remained constant. Thus, the determinant of product composition seems to lie exclusively in the structure of the primer DNA and the concentrations of the individual deoxyribonucleoside triphosphates do not play an important role here. It is clear from experiments cited earlier that elimination of one of the triphosphates entirely from the incubation mixture effectively stops DNA synthesis except under special circumstances. Therefore, although the *structure* of the product is not influenced by the concentration of the substrates, the *rate* of syn-

TABLE XIII

Purine and Pyrimidine Composition of Enzymatically Synthesized DNA[a]

|  | No. of analyses | A[b] | T[b] | G[b] | C[b] | $\frac{A+T}{G+C}$ | $\frac{A+G}{T+C}$ |
|---|---|---|---|---|---|---|---|
| *Myobacterium phlei* |  |  |  |  |  |  |  |
| Primer | 3 | 0.65 | 0.66 | 1.35 | 1.34 | 0.49(0.48–0.49) | 1.01(0.98–1.04) |
| Product | 3 | 0.66 | 0.80 | 1.17 | 1.34 | 0.59(0.57–0.63) | 0.85(0.78–0.88) |
| *Aerobacter aerogenes* |  |  |  |  |  |  |  |
| Primer | 1 | 0.90 | 0.90 | 1.10 | 1.10 | 0.82 | 1.00 |
| Product | 3 | 1.02 | 1.00 | 0.97 | 1.01 | 1.03(0.96–1.13) | 0.99(0.95–1.01) |
| *Escherichia coli* |  |  |  |  |  |  |  |
| Primer | 2 | 1.00 | 0.97 | 0.98 | 1.05 | 0.97(0.96–0.99) | 0.98(0.97–0.99) |
| Product | 2 | 1.04 | 1.00 | 0.97 | 0.98 | 1.02(0.96–1.07) | 1.01(0.96–1.06) |
| Calf thymus |  |  |  |  |  |  |  |
| Primer | 2 | 1.14 | 1.05 | 0.90 | 0.85 | 1.25(1.24–1.26) | 1.05(1.03–1.08) |
| Product | 6 | 1.19 | 1.19 | 0.81 | 0.83 | 1.46(1.22–1.67) | 0.99(0.82–1.04) |
| T2-bacteriophage |  |  |  |  |  |  |  |
| Primer | 2 | 1.31 | 1.32 | 0.67 | 0.70 | 1.92(1.86–1.97) | 0.98(0.95–1.01) |
| Product | 2 | 1.33 | 1.29 | 0.69 | 0.70 | 1.90(1.82–1.98) | 1.02(1.01–1.03) |
| Synthetic A-T copolymer[c] | 1 | 1.99 | 1.93 | <0.05 | <0.05 | <40 | 1.05 |

[a] From Lehman *et al.* (1958a).

[b] A, T, G, and C refer, respectively, to adenine, thymine, guanine, and cytosine except that C in the case of T2 bacteriophage primer refers to hydroxymethylcytosine. The values given represent averages of the number of analyses indicated. The figures in parentheses represent the range of values obtained.

[c] A copolymer of deoxyadenylate and thymidylate is formed by the "polymerase system in the absence of added DNA after a lag period of 3–6 hours. Once formed and isolated, this A-T polymer initiates the synthesis of new polymer without any time lag and, despite the presence of all four deoxynucleoside triphosphates in the incubation mixture, contains exclusively A and T.

TABLE XIV
EFFECT OF EXTENT OF NET SYNTHESIS ON COMPOSITION OF ENZYMATICALLY SYNTHESIZED DNA[a]

| | Expt. No. | Time (hours) | Increase in DNA (%) | T incorporated ($m\mu$moles) | C incorporated ($m\mu$moles) | T/C |
|---|---|---|---|---|---|---|
| *Mycobacterium phlei* DNA | | | | | | |
| Primer | | | | | | 0.49 |
| Product | 1 | | 2 | 0.047 | 0.11 | 0.42 |
| | 2 | 1.0 | 35 | 0.97 | 2.05 | 0.47 |
| | | 2.0 | 86 | 2.13 | 5.03 | 0.43 |
| | | 4.0 | 163 | 4.20 | 9.54 | 0.44 |
| | | 6.0 | — | 10.25 | 9.88 | 1.04 |
| | | 7.0 | — | 24.20 | 12.10 | 2.00 |
| Calf thymus DNA | | | | | | |
| Primer | | | | | | 1.24 |
| Product | 3 | | 8 | 0.40 | 0.28 | 1.43 |
| | 4 | 0.5 | 63 | 3.08 | 2.43 | 1.27 |
| | | 1.0 | 197 | 9.10 | 7.72 | 1.20 |
| | | 3.0 | 534 | 31.60 | 20.60 | 1.53 |
| *Aerobacter aerogenes* DNA | | | | | | |
| Primer | | | | | | 0.82 |
| Product | 5 | 0.5 | 18 | 0.74 | 0.93 | 0.79 |
| | | 1.0 | 56 | 2.59 | 2.89 | 0.91 |
| | | 3.0 | 344 | 15.00 | 17.75 | 0.84 |
| | | 4.0 | 406 | 16.62 | 21.00 | 0.80 |

[a] From Lehman *et al.* (1958a).

thesis of the DNA could be influenced by lowering the concentration of any one of the four deoxyribonucleoside triphosphates.

### 2. *Nearest Neighbor Analysis*

Although it is not yet possible to determine the sequence of nucleotides in DNA, it has been possible to examine the frequency of nucleotide pairs in products prepared from several different primers. The technique as developed by Josse *et al.* (1961) involves the following operations:

a. Synthesis of four identical products each one containing a different $P^{32}$-labeled deoxynucleotide. This is accomplished by setting up incubation mixtures containing the following deoxyribonucleoside triphosphates:

1. $dATP^{32} + dTTP + dGTP + dCTP$
2. $dATP + dTTP^{32} + dGTP + dCTP$
3. $dATP + dTTP + dGTP^{32} + dCTP$
4. $dATP + dTTP + dGTP + dCTP^{32}$

TABLE XV
Effect of Relative Substrate Concentrations on Composition of Synthetic DNA[a]

| Substrates (relative molar concentration) | | | | Net synthesis | Products | |
|---|---|---|---|---|---|---|
| dCTP | dGTP | dTTP | dATP | | $\frac{A+T}{G+C}$ | $\frac{A+G}{C+T}$ |
| 1.0 | 1.0 | 1.0 | 1.0 | 11× | 1.98 | 1.00 |
| 1.0 | 1.0 | 0.2 | 1.0 | 6× | 1.82 | 1.04 |
| 1.0 | 1.0 | 0.2 | 0.2 | 6× | 1.82 | 0.97 |
| T2 bacteriophage DNA as primer | | | | | 1.92 | 0.98 |

[a] From Lehman *et al.* (1958a).

b. Degradation of each of the products in such a way that the $P^{32}$ introduced into the DNA in the form of the specific deoxyribonucleotide is now part of the particular deoxyribonucleotide to which the labeled triphosphate was attached. This is accomplished by incubating the individual products with micrococcal DNase (Cunningham *et al.,* 1956), and splenic phosphodiesterase (Heppel *et al.,* 1953; Hilmoe, 1960).

c. Separation of the deoxyribonucleotides and estimation of their $P^{32}$ content. This is accomplished by paper electrophoresis (Markham and Smith, 1953).

For a detailed description of this technique, the reader is referred to the original paper of Josse *et al.* (1961). A brief description of a typical determination follows: Four different reaction mixtures containing one labeled and three unlabeled triphosphates are incubated and the reaction is terminated by the addition of carrier DNA and perchloric acid. Let us consider the product containing $dAMP^{32}$. If we examine Fig. 9 we see that $dATP^{32}$ (designated as Y-nucleotide) adds to the growing deoxyribonucleotide chain by forming a 3′–5′-phosphodiester bridge with the terminal deoxynucleotide (designated X-nucleotide). This will occur at specific points in the newly synthesized DNA and these positions are presumably dictated by base-pairing with the template DNA. The product is washed free of acid-soluble unreacted deoxyribonucleoside triphosphates and degraded enzymatically as shown in Fig. 9. Using the same designation as for synthesis, it can be seen that since the enzymes specifically hydrolyze the phosphodiester bridges between the 5′-carbon and the phosphorus, the $P^{32}$ from the incorporated $dAMP^{32}$ (Y) now ends up on the 3′-position of the neighboring deoxyribonucleotide (X). These 3′-deoxyribonucleotides are separated and analyzed for $P^{32}$ content and the fraction of each of the four possible dinucleotide pairs containing dAMP as the 5′-deoxyribonucleotide can be calculated. This value, multiplied by the molar proportion of dAMP in the particular nucleic acid studied (this may be determined independently or it can be derived from the nearest neighbor data) gives the fraction of the particular dinucleotide in the DNA. The procedure is repeated for the other three reaction mixtures and the frequencies of all possible dinucleotides may be determined.

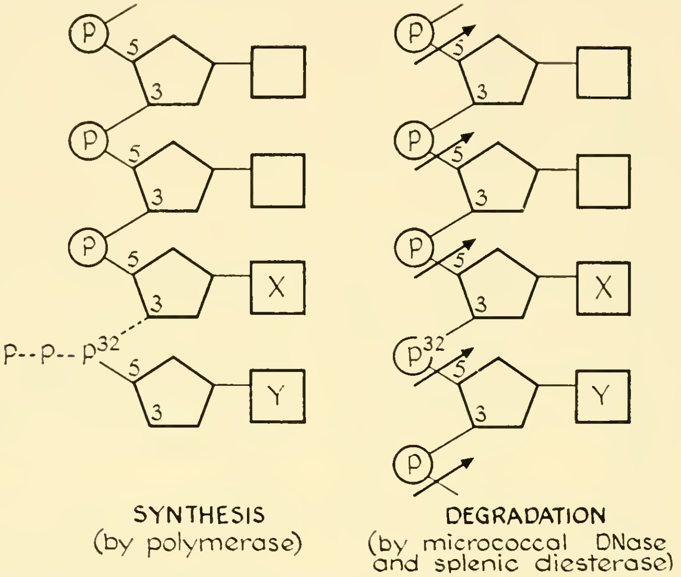


Fig. 9. Synthesis of a $P^{32}$-labeled DNA chain and its subsequent enzymatic degradation to 3′-deoxyribonucleotides. The arrows indicate the linkages cleaved by micrococcal DNase and calf spleen phosphodiesterase, yielding a digest composed exclusively of 3′-deoxyribonucleotides. (From Josse *et al.*, 1961.)

Table XVI illustrates the data obtained from a typical nearest neighbor analysis. In this experiment, *Mycobacterium phlei* DNA was used as primer and the actual radioactivity present in each of the fractions is recorded. In the columns labeled "sequence" are reported the dinucleotides represented by the incorporated label. In the columns labeled "fraction" are the values calculated as the fraction of the total counts in a particular reaction digest found in the particular 3′-deoxyribonucleotides. When these fractions are multiplied by the base-incorporation factors (mole fraction of the particular base in the given DNA) the nearest neighbor frequencies recorded in Table XVII result. Earlier work showing that the base composition of the newly synthesized DNA reflected that of the primer was consistent with the scheme in Fig. 10 in which each strand of the proposed Watson-Crick double helix acts as

TABLE XVI

Radioactivity Measurements in Experiments with *Mycobacterium phlei* DNA as Primer[a]

| Isolated 3′-deoxyribonucleotide | Labeled triphosphate | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Reaction No. 1 (dATP$^{32}$) | | | Reaction No. 2 (dTTP$^{32}$) | | | Reaction No. 3 (dGTP$^{32}$) | | | Reaction No. 4 (dCTP$^{32}$) | | |
| | Sequence | cpm | Fraction | Sequence | cpm | Fraction | Sequence | cpm | Fraction | Sequence | cpm | Fraction |
| Tp | TpA | 873 | 0.075 | TpT | 1,665 | 0.157 | TpG | 3,490 | 0.187 | TpC | 4,130 | 0.182 |
| Ap | ApA | 1,170 | 0.146 | ApT | 2,065 | 0.194 | ApG | 2,500 | 0.134 | ApC | 4,300 | 0.189 |
| Cp | CpA | 4,430 | 0.378 | CpT | 2,980 | 0.279 | CpG | 7,730 | 0.414 | CpC | 6,070 | 0.268 |
| Gp | GpA | 4,690 | 0.401 | GpT | 3,945 | 0.370 | GpG | 4,960 | 0.265 | GpC | 8,200 | 0.361 |
| Sum | | 11,703 | 1.000 | | 10,655 | 1.000 | | 18,680 | 1.000 | | 22,700 | 1.000 |

[a] From Josse *et al.* (1961).

TABLE XVII

NEAREST NEIGHBOR FREQUENCIES OF *Mycobacterium phlei* DNA[a,b]

| Reaction No. | Labeled triphosphate | Isolated 3'-deoxyribonucleotide[c] | | | |
|---|---|---|---|---|---|
| | | Tp | Ap | Cp | Gp |
| 1 | dATP$^{32}$ | a<br>TpA<br>0.012 | b<br>ApA<br>0.024<br>I | c<br>CpA<br>0.063<br>II | d<br>GpA<br>0.056<br>III |
| 2 | dTTP$^{32}$ | b<br>TpT<br>0.026<br>I | a<br>ApT<br>0.031 | d<br>CpT<br>0.045<br>IV | c<br>GpT<br>0.060<br>V |
| 3 | dGTP$^{32}$ | e<br>TpG<br>0.063<br>II | f<br>ApG<br>0.045<br>IV | g<br>CpG<br>0.139 | h<br>GpG<br>0.090<br>VI |
| 4 | dCTP$^{32}$ | f<br>0.061<br>III | e<br>0.064<br>V | h<br>0.090<br>VI | g<br>0.122 |
| | Sums | 0.162 | 0.164 | 0.337 | 0.337 |

[a] From Josse *et al.* (1961).

[b] Chemical analysis of the base composition of the primer DNA indicated molar proportions of thymine, adenine, cytosine, and guanine of 0.165, 0.162, 0.335, and 0.338, respectively.

[c] Identical Roman numerals designate those sequence frequencies which should be equivalent in a Watson and Crick DNA model with strands of opposite polarity; identical lower case letters designate sequence frequencies which should be equivalent in a model with strands of similar polarity.

a template for the synthesis of a daughter strand with complementary bases. The nearest neighbor frequencies of Table XVII provide independent evidence for this scheme. ApA and TpT sequences are equal as are GpG and CpC. In examining the other sequences, it is important to realize that the predicted pairings of dinucleotides will depend on whether the two chains of the double helix run in the same or opposite directions, that is, have similar or opposite polarity. This difference is made clear in Fig. 11 where a short segment of the replicating chain is depicted. The expected pairing of dinucleotide-sequences as a consequence of similar and opposite polarity in the nucleic acid chains are shown, and it is obvious that decidedly different pairings would be expected depending on which of the two possibilities obtained. In Table

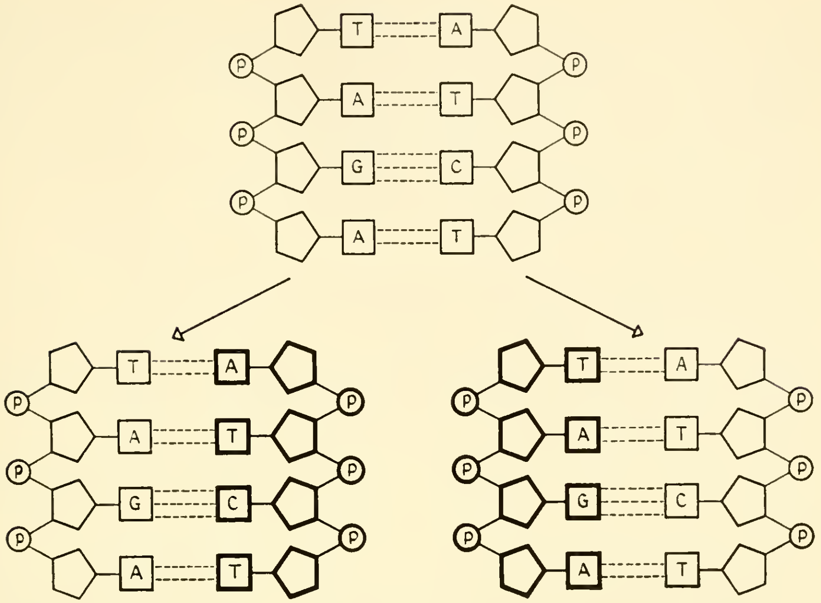


FIG. 10. Proposed scheme of replication of a Watson and Crick DNA model. Bold-lined polynucleotide chains of the two daughter molecules represent newly synthesized strands. (From Josse *et al.,* 1961.)

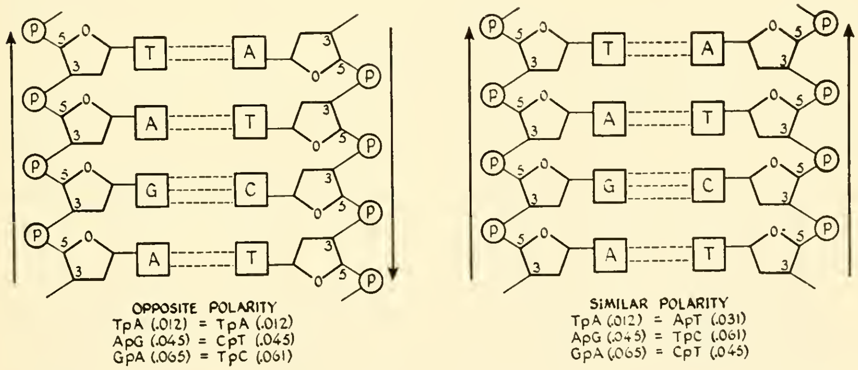

P
5
O
3
T
A
A
T
G
C
A
T
OPPOSITE POLARITY
TpA (.012) = TpA (.012)
ApG (.045) = CpT (.045)
GpA (.065) = TpC (.061)
SIMILAR POLARITY
TpA (.012) = ApT (.031)
ApG (.045) = TpC (.061)
GpA (.065) = CpT (.045)

FIG. 11. Contrast of a Watson and Crick DNA model with strands of opposite polarity with a model with strands of similar polarity. The predicted matching nearest neighbor sequence frequencies are different. Values in parentheses are sequence frequencies from the experiments with *M. phlei* DNA. The strands represented here are the newly synthesized strands of Fig. 10; for ease of comparison, they are aligned as if they were complementary strands of the same double helix. (From Josse *et al.,* 1961.)

XVII, the expected pairings based on strands of opposite polarity are designated by Roman numerals whereas the pairings based on similar polarity are designated by lower case letters. It is clear that the fit is in excellent agreement with the predictions based on opposite polarity and not with the other possibility. This represents an elegant proof of the original proposal that the DNA molecule is composed of strands of opposite polarity (Watson and Crick, 1953a).

This technique was extended to other DNA types with the following observations:

1. Nearest neighbor sequence frequencies are the same in products primed with heated or unheated DNA.

2. The nearest neighbor sequence frequencies of DNA from λ and λ-dg (temperate bacteriophages) were very similar to those of the host, whereas the frequencies in T2, T4, and T6 (virulent bacteriophages) were markedly different from those of the host.

3. The dinucleotide frequencies of DNA from most sources were unique to the species and did not conform to a calculated random distribution of the bases.

4. The dinucleotide frequencies of the DNA preparations from several tissues of the same organism showed no significant variation, nor did the DNA preparations from normal and neoplastic tissues of the same organism (Swartz *et al.*, 1962).

5. The distribution of sequence frequencies in 16 DNA samples analyzed conform to a mechanism for DNA replication involving pairing of adenine to thymine and guanine to cytosine and are in keeping with the opposite polarity of the DNA strands.

6. The nearest neighbor sequence frequencies of a product primed with native calf thymus DNA or primed with newly synthesized "calf thymus" DNA (containing less than 5% of the original native DNA) were the same. From this it may be inferred (though not proved) that the sequence frequencies in the product are indicative of the sequence frequencies in the native DNA.

An interesting application of the nearest neighbor analysis is the recent investigation (Swartz *et al.*, 1962) of whether both strands of the double-strand form of DNA from the small bacteriophage φX174 (Sinsheimer, 1959) can serve as templates in DNA replication. When the single-stranded form of φX174 is used to prime a reaction in which each strand of the original DNA replicates once, the composition of the newly formed strand should be complementary to the original. An experiment was devised in which DNA synthesis was allowed to proceed until 20% of the original DNA was replicated, thus minimizing the chance of multiple rounds of replication. The product of this reaction was analyzed

TABLE XVIII
Composition of Products after Limited and Extensive Replication of $\phi$X DNA[a]

| | | Composition determined by nearest neighbor analysis | | | | |
|---|---|---|---|---|---|---|
| | | 20% synthesis | | 600% synthesis | | |
| Base | Composition determined by chemical analysis | Predicted from chemical analysis | Observed | Predicted[b] from chemical analysis | Predicted[b] from 20% synthesis | Observed |
| A | 0.246 | 0.328 | 0.310 | 0.287 | 0.276 | 0.271 |
| T | 0.328 | 0.246 | 0.242 | 0.287 | 0.276 | 0.293 |
| G | 0.242 | 0.185 | 0.202 | 0.214 | 0.224 | 0.213 |
| C | 0.185 | 0.242 | 0.246 | 0.214 | 0.224 | 0.224 |

[a] From Swartz *et al.* (1962).

[b] Based on unlimited replication.

for base composition by the nearest neighbor technique (for details of the use of nearest neighbor analysis for determining base composition see Josse *et al.*, 1961) and the results reported in Table XVIII show that the observed values agree very well with the predicted values. Based on this "limited" synthesis, the base composition of product formed after

TABLE XIX
Nearest Neighbor Frequencies[a] of $\phi$X DNA in Limited and in Extensive Replication[b]

| | Limited replication (20%) | Extensive replication (600%) | |
|---|---|---|---|
| Nearest neighbor sequence | Observed | Predicted from limited replication[c] | Observed |
| ApA, TpT | 0.101, 0.069 | 0.085, 0.085 | 0.085, 0.099 |
| CpA, TpG | 0.096, 0.048 | 0.072, 0.072 | 0.070, 0.070 |
| GpA, TpC | 0.054, 0.064 | 0.059, 0.059 | 0.058, 0.065 |
| CpT, ApC | 0.052, 0.069 | 0.061, 0.061 | 0.064, 0.058 |
| GpT, ApC | 0.047, 0.068 | 0.057, 0.057 | 0.053, 0.053 |
| GpG, CpC | 0.040, 0.053 | 0.046, 0.046 | 0.041, 0.045 |
| TpA | 0.061 | 0.061 | 0.059 |
| ApT | 0.072 | 0.072 | 0.075 |
| CpG | 0.045 | 0.045 | 0.045 |
| GpC | 0.061 | 0.061 | 0.061 |

[a] Expressed as decimal proportions of 1.000.

[b] From Swartz *et al.* (1962).

[c] Based on unlimited replication.

extensive synthesis (6-fold) was calculated and compared to the experimental findings. Again there is good agreement between the observed and predicted values. The calculation was based on the assumption that both strands replicate equally and this appears to be borne out by the data. The results were extended by subjecting the products of "limited" and "extensive" synthesis to nearest neighbor analysis. Based on the assumption that both strands replicate equally well, values were calculated for nearest neighbor frequencies in the DNA from "extensive" synthesis and the data in Table XIX support this contention very well. These elegant experiments further support the proposed scheme for replication (Fig. 10) in which both strands of the primer DNA act as templates through which specific base-pairing is accomplished.

A few of the applications of nearest neighbor analysis have been mentioned here. It is clear that until a new technique is devised for determining longer sequences in the DNA molecule, the nearest neighbor approach will provide the most revealing information about the fine structure of DNA.

### G. SYNTHESIS OF DEOXYADENYLATE-DEOXYTHYMIDYLATE AND DEOXYGUANYLATE-DEOXYCYTIDYLATE POLYMERS

(Schachman *et al.*, 1960)

We have already seen an exception to our generalization that all four deoxyribonucleoside triphosphates are required for DNA synthesis in that single deoxyribonucleotides can add to the ends of DNA chains. In this section, two reactions catalyzed by *E. coli* polymerase, which not only do not require all four triphosphates, but in addition may proceed in the absence of primer, will be discussed.

#### 1. *The d-AT Polymer*

When *E. coli* polymerase is incubated with all four triphosphates in the absence of a primer there is an extensive lag period, 3 to 4 hours, in which no detectable reaction occurs. This period of inactivity is followed by a rapid synthesis of a high molecular weight polymer of deoxyadenylate and deoxythymidylate. The kinetics of the reaction are shown in Fig. 12, where three independent techniques were used to follow the reaction. In Expt. 1, plots of the rate of incorporation of $P^{32}$ from $dATP^{32}$ and the rate of increase in viscosity are superimposable. In Expt. 2, the hypochromicity curve is the mirror image of the $P^{32}$ incorporation. In Fig. 13, the kinetics of DNA synthesis in the complete system is compared to the synthesis of the d-AT polymer. The complete system contained, beside polymerase and calf thymus DNA, all 4 triphosphates. Polymer formation as measured by viscosity increase starts

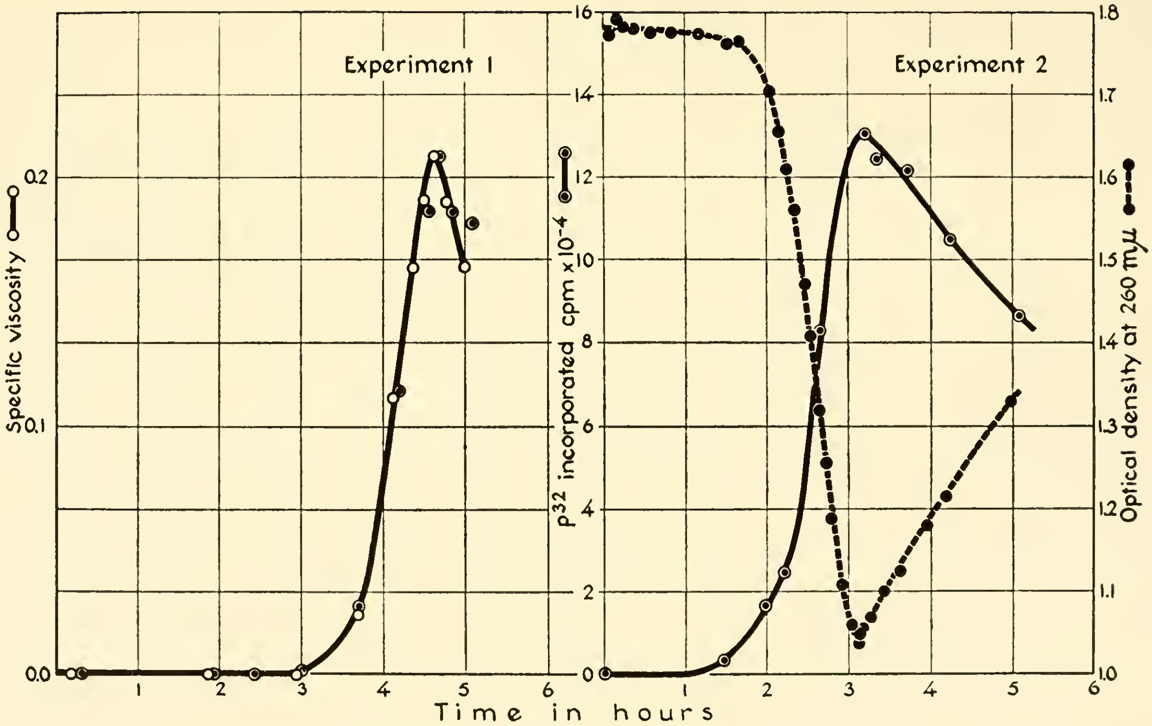


FIG. 12. Course of synthesis of d-AT copolymer as measured by viscometry, spectrophotometry, and radioisotope incorporation. The reaction mixture contained, in a volume of 1 ml: potassium phosphate buffer, pH 7.4, 60 μmoles; $MgCl_2$, 6.0 μmoles; dTTP, 0.3 μmole; $dATP^{32}$ ($5.3 \times 10^5$ cpm per μmole), 0.32 μmole; enzyme, about 3 units. No DNA was present. (From Schachman *et al.*, 1960.)

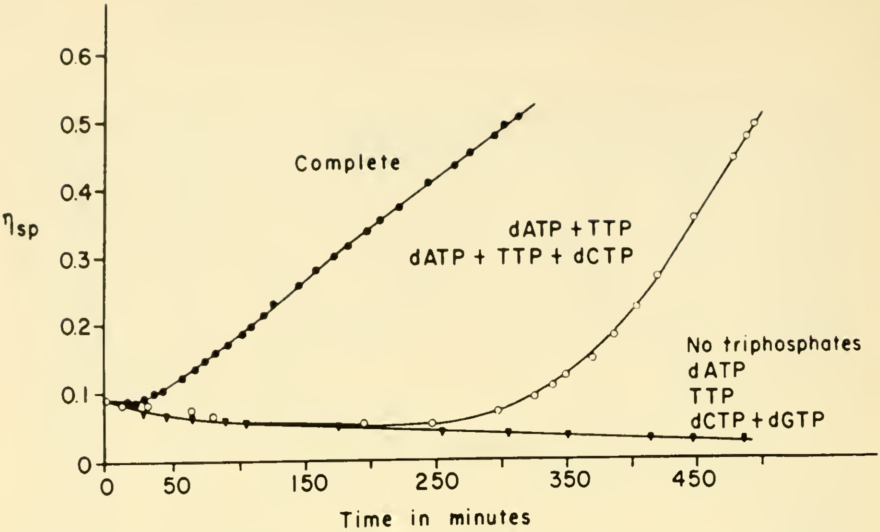


Fig. 13. Synthesis of DNA and deoxyadenylate-thymidylate copolymer, with thymus DNA as primer. (From Lehman, 1959.)

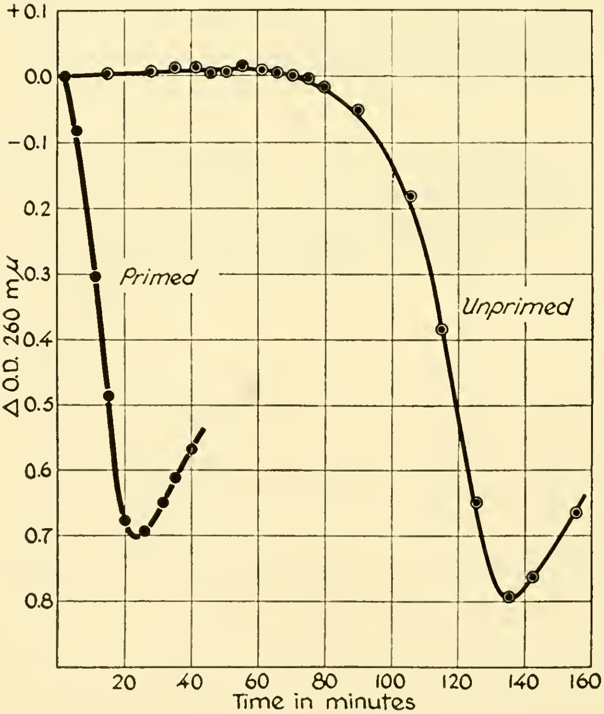


Fig. 14. Comparison of primed *versus* unprimed synthesis of d-AT copolymer. (From Schachman *et al.*, 1960.)

with no appreciable lag and continues linearly. When dGTP or both dGTP and dCTP are omitted from the incubation, there is a 4-hour lag followed by a period of rapid synthesis. If no triphosphates are added to the incubation mixture, or just those indicated in the last curve in the figure, no synthesis occurs. The lag period can be reduced somewhat by raising the enzyme concentration; but it can be eliminated completely by priming the reaction with synthetic d-AT polymer (Fig. 14). Only d-AT polymer itself is an effective primer in the reaction, and other DNA preparations have no effect in reducing the lag period.

When the "limited" reaction (Section IV,C) is investigated, using d-AT as the primer, small but significant quantities of deoxyadenylate or deoxythymidylate are added to the ends of the polymer. It is of interest that neither dCMP nor dGMP can add to the ends of the d-AT polymer, suggesting that even in the limited reaction specific base-pairing is required. This requirement for specific base-pairing favors the "repair" hypothesis suggested as one of the possible explanations of the "limited" reaction (Section IV,C).

Analysis of the d-AT polymer shows it to be a rodlike organized structure with a molecular weight ranging between $10^6$ and $10^7$. It is composed exclusively of equal proportions of deoxyadenylate and deoxythymidylate residues, and an interesting feature is that these residues are arranged in a strictly alternating sequence. This is demonstrated in Table XX where the results of a nearest neighbor analysis show that each deoxythymidylate residue is linked by a phosphodiester bridge to a deoxyadenylate residue and vice versa. This regularly alternating sequence has been used to explain the observation that the d-AT co-

TABLE XX

HYDROLYSIS OF d-AT COPOLYMER TO 3′-DEOXYRIBONUCLEOTIDES[a,b]

| | | Products of hydrolysis | | | |
|---|---|---|---|---|---|
| | | 3′-Deoxyadenylate | | 3′-Deoxythymidylate | |
| Expt. No. | Substrates for preparation of d-AT copolymer | cpm | cpm in polymers (%) | cpm | cpm in polymers (%) |
| 1 | dAPPP and $dTP^{32}PP$ | 9670 | 100 | 23 | <0.5 |
| 2 | $dAP^{32}PP$ and dTPPP | 46 | <0.5 | 8510 | 100 |

[a] From Schachman *et al.* (1960).

[b] d-AT copolymer was hydrolyzed to 3′-deoxyribonucleotides by micrococcal DNase and splenic phosphodiesterase until more than 96% of the $P^{32}$ was released in the form of deoxyribonucleoside monophosphates.

polymer does not show the hysteresis typical of DNA preparations after heating and cooling (Fig. 15). It has been suggested that the regularly alternating sequence of deoxyadenylate and deoxythymidylate in each strand increases the opportunity for correct alignment of the bases in opposite strands during the cooling period (Marmur and Doty, 1959; Schachman *et al.*, 1960).

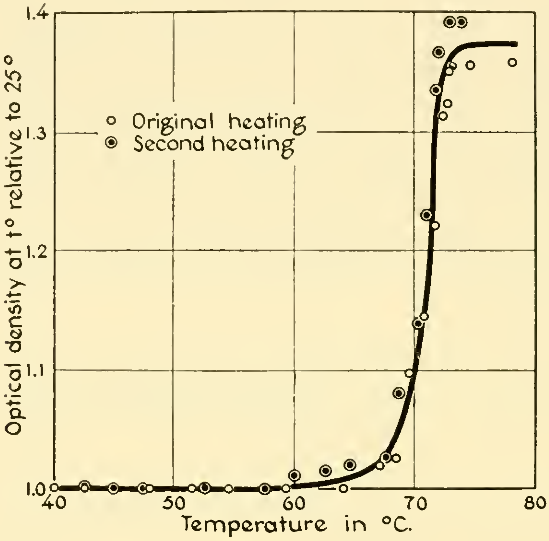


FIG. 15. Melting of d-AT copolymer. Plotted on the ordinate is the relative absorbancy at 260 mμ at $t°$ compared to the value at 25°C. The abscissa gives the temperature. The first heating cycle is indicated by open circles, and the solution after cooling in the stoppered Beckman cell was heated a second time, giving the data indicated. (From Schachman *et al.*, 1960.)

Attempts to elucidate the mechanism of synthesis have so far met with little success. For example, in the unprimed synthesis of d-AT polymer, at any stage during the synthesis only two types of molecules are observed. One type has a molecular weight in the order of $10^3$ and is obviously made up of the original triphosphate molecules, and the other has a molecular weight in the order of $10^6$ and must represent completed molecules. No intermediates have been found between these two classes of molecules, as indicated in Table XXI where it can be seen that the sedimentation coefficients and reduced viscosities are constant throughout the early and later phases of the synthesis.

Recently, Sueoka (1961) reported the presence of a minor component in the DNA of crab testes (*Cancer borealis*) which had an unusually low density suggesting a very low content of cytosine and guanine. A sample of this material was used to prime a reaction using *E. coli* polymerase and the product was analyzed for nearest neighbor sequence frequencies (Swartz *et al.*, 1962). Over 95% of the product had a sequence made up of alternating deoxyadenylate and deoxythymidylate residues. This naturally occurring d-AT polymer differs from the unprimed d-AT polymer produced by *E. coli* polymerase in that it does have a small but significant content of deoxyguanylate and deoxycytidylate residues accounting for less than 5% of the total deoxyribonucleotides.

TABLE XXI

RELATIONSHIP BETWEEN EXTENT OF REACTION AND SIZE OF PRODUCT[a,b]

| | Extent of reaction | | Properties of product | |
|---|---|---|---|---|
| Expt. No. | Viscosity (%) | Hypochromicity (%) | Sedimentation coefficient (S) | Reduced viscosity (gm/100 ml)$^{-1}$ |
| 1 | 8 | 5 | 29 | 26 |
| | 37 | 46 | 28 | 26 |
| | 100 | 100 | 27 | 24 |
| 2 | 36 | 28 | 19 | 23 |
| | 80 | 91 | 19 | 25 |
| | 100 | 100 | 20 | 26 |

[a] From Schachman *et al.* (1960).

[b] Reaction mixtures, 6 ml, were incubated directly in viscometers; at selected times after the viscosity of the solution had been measured, aliquots were removed, and the enzymatic action terminated by the addition of NaCl-citrate solution to give a concentration of 0.2 *M* NaCl–0.1 *M* sodium citrate. Absorbancy measurements of these samples as compared to the original gave a measure of the extent of the reaction (hypochromicity). The samples were then dialyzed, and the reduced viscosity, $\eta_{sp}/C$, in (gm/100 ml)$^{-1}$ and sedimentation coefficient were measured in 0.02 *M* sodium citrate.

### 2. *The d-GC Polymer*

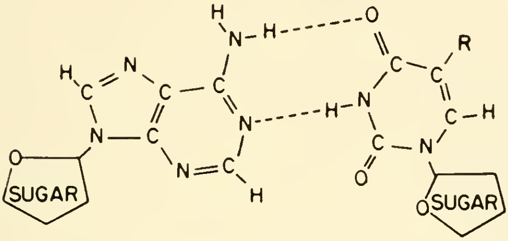
Much less information is available about an interesting polymer composed exclusively of deoxycytidylate and deoxyguanylate residues. This material is similar to the d-AT copolymer in that it is composed of only two deoxyribonucleotides and is formed after a considerable lag period in an unprimed reaction catalyzed by purified *E. coli* polymerase. However, this product is distinctly different in one important respect. Whereas d-AT is a copolymer of alternating residues of deoxyadenylate and deoxythymidylate, the d-GC polymer is made up of separate strands containing either deoxyguanylate or deoxycytidylate. This was shown clearly by Josse *et al.* (1961) by nearest neighbor analysis. Within the limits of the methods, deoxycytidylate was only found next to deoxycytidylate. No phosphodiester linkages were evident between deoxyguanylate and deoxycytidylate. Besides this difference, d-GC polymer is distinct from d-AT in having unequal proportions of deoxycytidylate and deoxyguanylate. Usually, the molar ratio of dGMP to dCMP is greater than 1.

### H. SUBSTITUTION OF ANALOGS

An analog is any purine or pyrimidine base not normally found in the DNA of a particular organism. For example, although 5-methylcytosine is a normal constituents of wheat germ DNA, it would be considered an analog if incorporated into *E. coli* DNA, because it is not normally detected there. Several analogs have been synthesized as the deoxyribonucleoside triphosphates and tested for incorporation into DNA (Bessman *et al.*, 1958b). When primed with *E. coli* DNA, the following analogs support DNA synthesis: uracil, hypoxanthine, 5-bromouracil, 5-bromocytosine, and 5-methylcytosine. Xanthine was not incorporated. Recent experiments from Kornberg's laboratory (personal communication) have shown that 5-fluorocytosine and 5-fluorouracil are incorporated into DNA whereas *N*-methyl 5-fluorocytosine is not. 5-Hydroxymethylcytosine, the pyrimidine base found exclusively in the T-even bacteriophages (Wyatt and Cohen, 1953), is also incorporated into DNA by *E. coli* polymerase primed with DNA prepared from several different sources including calf thymus, *E. coli* and phage $\phi$X174 (Kornberg *et al.*, 1959).

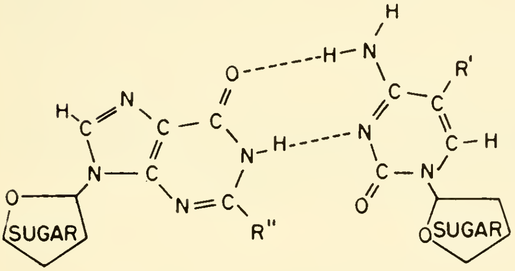
According to the restrictions imposed by the base-pairing rules of the Watson-Crick model (1953b), these base analogs should be incorporated into specific positions in the DNA molecule. A diagram of these relationships is shown in Fig. 16. The specificity in replacing a given base by an analog was tested by determining which deoxyribonucleoside triphosphate could be omitted from a reaction mixture in the presence of a given analog and still support DNA synthesis. For example, in order to find out which substrate could be replaced by deoxyuridine triphosphate, four reactions were set up, each lacking one of the normal triphosphates. DNA synthesis proceeded only in the reaction mixture lacking dTTP, suggesting that dUTP was replacing the dTTP requirement. Table XXII contains the results of several such experiments. It can be seen that in each case the given analog substitutes for the purine or pyrimidine base predicted by the Watson-Crick model. Table XXIII shows direct evidence for the incorporation of deoxyuridylate into DNA. The results are consistent with a specific substitution of deoxyuridylate for deoxythymidylate since the data indicate that these two nucleotides compete with each other during the synthesis. Results of a nearest neighbor analysis on a product synthesized with $dUTP^{32}$ showed that deoxyuridylate was linked to each of the other deoxyribonucleotides.

Hydrogen Bonding of Adenine to Thymine



| R | |
|---|---|
| $-CH_3$ | Thymine |
| $-H$ | Uracil |
| $-Br$ | Bromuracil |

Hydrogen Bonding of Guanine to Cytosine



| R' | |
|---|---|
| $-H$ | Cytosine |
| $-CH_3$ | Methylcytosine |
| $-Br$ | Bromcytosine |

| R" | |
|---|---|
| $-NH_2$ | Guanine |
| $-H$ | Hypoxanthine |

FIG. 16. Base-pairing relationships of normal constituents and analogs as predicted from the Watson and Crick model for DNA. (From Bessman *et al.*, 1958b.)

An interesting study of the effect of 5-bromouracil substitution on replication of DNA has recently been reported (Trautner *et al.*, 1962). The d-AT polymer was used to test the frequency with which mistakes in base-pairing occur during DNA synthesis by noting the frequency with which dGMP was incorporated into the d-AT product. Highly radioactive $dGTP^{32}$ (specific activity $= 1 \times 10^8$ cpm/$\mu$mole) provided the basis for an assay that was capable of detecting the incorporation of

### TABLE XXII
### Replacement of Natural Bases by Analogs in Enzymatic Synthesis of DNA[a]

| Expt. No. | Control value[b] (mμmoles) | Base analog used | Natural base omitted: Thymine | Adenine | Guanine | Cytosine |
|---|---|---|---|---|---|---|
| | | | (% of control)[c] | | | |
| 1 | 0.50 | Uracil | *54* | 4 | 6 | |
| 1a | 0.88 | Uracil | | | | 3 |
| 2 | 0.43 | 5-Bromouracil | *97* | 2 | 4 | |
| 2a | 0.42 | 5-Bromouracil | | | | 4 |
| 3 | 0.51 | 5-Bromocytosine | | 4 | 4 | *118* |
| 3a | 0.40 | 5-Bromocytosine | 4 | | | |
| 4 | 0.58 | 5-Methylcytosine | | 2 | 3 | *185* |
| 4a | 0.52 | 5-Methylcytosine | 2 | | | |
| 5 | 0.37 | Hypoxanthine | | 3 | *25* | 5 |
| 5a | 0.27 | Hypoxanthine | 4 | | | |

[a] From Bessman *et al.* (1958b).

[b] Control values are mμmoles of radioactive deoxynucleotide incorporated into DNA in the absence of analog. Incubation mixtures contained in 0.3 ml, 5 mμmoles each of TTP, dATP, dCTP, and dGTP; 2 μmoles of $MgCl_2$; 20 μmoles of potassium phosphate (pH 7.4); 10 μg of calf-thymus DNA; and 1 μg of enzyme fraction VII-R. Experiments were performed at 37°C for 30 minutes. Labeled substrates were: $dCP^{32}PP$ in Expts. 1, 2, and 5a; $TP^{32}PP$ in Expts. 1a, 3, 4, and 5; and $dGP^{32}PP$ in Expts. 2a, 3a, and 4a.

[c] The percentage value represents the fraction of the labeled substrate incorporated when the analog (5 mμmoles) was used instead of a natural base. All bases, natural or analog, were supplied as the deoxynucleoside triphosphates. Values of 5% or below are near the limit of detectability and are of questionable significance.

### TABLE XXIII
### Incorporation of $P^{32}$-Deoxyuridylate into DNA[a]

| Substrates[b] | $P^{32}$-deoxynucleotide incorporated (mμmoles) |
|---|---|
| $\overset{*}{U} + T + C + A + G$ | 0.14 |
| Omit T | 0.38 |
| Omit C | 0.01 |
| Omit A | 0.01 |
| Omit G | 0.01 |
| $U + \overset{*}{T} + C + A + G$ | 0.37 |
| Omit U | 0.77 |

[a] From Bessman *et al.* (1958b).

[b] U = dUTP; $\overset{*}{U}$ = $dUP^{32}PP$; T = TTP; $\overset{*}{T}$ = $TP^{32}PP$; C = dCTP; A = dATP; G = dGTP.

1 dGMP residue per $10^5$ nucleotide units. Within the limits of the assay, no significant incorporation of dGMP occurred in a reaction mixture made up of d-AT primer, $dGTP^{32}$, dATP, dTTP, $Mg^{++}$ and *E. coli* polymerase. However, when a polymer composed of dAMP and 5-bromodeoxyuridylate (d-A$\overline{\mathrm{BU}}$) was used in place of d-AT as the primer, there was a significant incorporation of dGMP. For example, in one experiment less than 0.10 $\mu\mu$moles of $dGMP^{32}$ was incorporated using d-AT as a primer, whereas 2.82 $\mu\mu$moles of $dGMP^{32}$ was incorporated using d-A$\overline{\mathrm{BU}}$ as primer.

Mistakes in base-pairing could result from tautomeric shifts in the bases during replication (Watson and Crick, 1953b). The specific mutagenic effect of 5-bromouracil which is incorporated into DNA in place of thymine (Dunn and Smith, 1954) has been attributed to its facilitated tautomerization to a form which could pair with guanine instead of adenine (Watson and Crick 1953b; Freese, 1959). If guanine only substituted for adenine in the d-A$\overline{\mathrm{BU}}$ polymer described above, a nearest neighbor analysis of a product prepared $dGMP^{32}$ as the labeled deoxyribonucleotide should reveal only one labeled 3′-deoxyribonucleotide, that is, 5-bromodeoxyuridylate. When Trautner *et al.* (1962) analyzed the product they found that all three possible 3′-deoxyribonucleotides were labeled (Table XXIV). These data show there are just as many

TABLE XXIV
NEAREST NEIGHBORS OF THE GUANINE RESIDUES INCORPORATED INTO THE d-A$\overline{\mathrm{BU}}$ POLYMER[a]

| Expt. No. | Radioactivity in the isolated 3′-deoxynucleotides | | | | | |
|---|---|---|---|---|---|---|
| | BU | | A | | G | |
| | cpm | % of total | cpm | % of total | cpm | % of total |
| 1 | 122 | 36.3 | 56 | 16.7 | 158 | 47.0 |
| 2 | 227 | 45.3 | 83 | 16.7 | 190 | 38.0 |

[a] From Trautner *et al.* (1962).

GG pairs as G$\overline{\mathrm{BU}}$ pairs in the product and about one-half as many GA pairs. This would suggest that the mutagenic effect of 5-bromouracil is not entirely explained on the basis of a mistake in pairing with adenylate instead of guanylate, but that the analog exerts a more complicated effect. The authors are careful in pointing out that these results *in vitro* with the d-A$\overline{\mathrm{BU}}$ polymer do not necessarily represent the *in vivo* effect of 5-bromouracil on whole DNA. Also, the argument outlined above is predicted on the basis that the d-AT (and hence d-A$\overline{\mathrm{BU}}$) polymer is composed of 100% alternating deoxyadenylate-deoxythymidylate residues.

Another interesting experiment of Trautner *et al.* (1962) was the preparation of a product from a reaction mixture primed with *E. coli* DNA in which 100% of the dTMP residues were replaced by 5-bromodeoxyuridylate (Baldwin *et al.,* 1961). A nearest neighbor analysis of this DNA revealed that its dinucleotide sequences were indistinguishable (±2%) from those in normal *E. coli* DNA. These results are not in keeping with a report of Shapiro and Chargaff (1960), based on the analysis of an acid-hydrolyzed product that marked changes in sequences in *E. coli* DNA result from substitution of 5-bromouracil.

The general conclusion that may be drawn from the studies on the incorporation of base analogs into DNA is that, *in vitro,* purine and pyrimidine bases supplied as deoxyribonucleoside triphosphates usually are incorporated into DNA at positions predicted by the base-pairing relationships of the Watson-Crick model. This does not mean that these same analogs supplied as free bases, deoxyribonucleosides, or deoxyribonucleotides will be incorporated into DNA *in vivo* since these compounds may be excluded by being blocked at any one of the several steps leading to the deoxyribonucleoside triphosphates (Kalle and Gots, 1961; Friedkin and Kornberg, 1957; Bessman *et al.*, 1958b). Another interesting method of exclusion is based on the specific hydrolysis of the deoxynucleoside triphosphates by enzymes such as dCTPase (Zimmerman and Kornberg, 1961; Koerner *et al.,* 1960) and dUTPase (Bertani *et al.,* 1961; Greenberg and Somerville, 1962a). These two enzymes are presumably involved in keeping cytosine out of the DNA of T-even bacteriophages and uracil out of *E. coli* DNA, respectively.

## V. SYNTHESIS OF DNA BY CALF THYMUS POLYMERASE

There have been numerous reports of the synthesis of DNA *in vitro* catalyzed by extracts prepared from higher organisms. With the exception of very few, these have been done with unfractionated preparations which contain several activities related and unrelated to the direct reaction. These investigations, although important in establishing an over-all picture of the fate of labeled nucleic acid precursors in cell homogenates cannot provide the detailed information necessary to define the pathways and mechanisms of DNA synthesis. Bollum (1959a,b, 1960a,b) has investigated the synthesis of DNA using a partially purified enzyme from calf thymus glands. His experiments will be discussed as representative of the properties of DNA synthesis in animals although the generalization may not be applicable to all systems. No obvious differences between the synthesis of DNA in the calf thymus system and the less purified regenerating rat liver system described by Bollum and Potter (1958) and Bollum (1958) are evident and therefore remarks will be confined to the former.

At the outset it may be stated that the reaction catalyzed by calf thymus polymerase (purified about 50-fold) is very similar to the one from *E. coli*. The requirement for all four deoxyribonucleoside triphosphates, Mg ions, and DNA is evident in Table XXV. Omission of any of these constituents reduces the rate of incorporation. The reaction is inhibited by high concentrations of pyrophosphate and it can be shown that DNA synthesis is reversible in this system by measuring the formation of deoxyribonucleoside triphosphates in the presence of polymerase, DNA, Mg ions, and pyrophosphate. With the partially purified enzyme, net synthesis of DNA can be demonstrated if heated DNA is used as the primer. The products formed resemble the primer with respect to adenine-cytosine or thymine-cytosine content as demonstrated in Table XXVI. Again, the failure of native DNA to prime the reaction is demonstrated in this table.

TABLE XXV
REQUIREMENTS FOR THYMIDINE TRIPHOSPHATE INCORPORATION WITH CALF THYMUS POLYMERASE[a]

| Reaction mixture | $H^3$-dTMP incorporated into DNA ($\mu\mu$moles) |
|---|---|
| Complete[b] | 100 |
| Minus dATP | 9 |
| Minus dGTP | 13 |
| Minus dCTP | 28 |
| Minus dATP, dGTP, dCTP | 2 |
| Minus $Mg^{++}$ | 0 |
| Minus DNA | 0 |

[a] From Bollum (1960a).

[b] The complete reaction mixture contained the following concentrations: deoxyribonucleoside triphosphates, 25 $\mu$moles in each ($H^3$-TTP containing 6700 cpm per m$\mu$mole); $MgCl_2$, 10 mmoles; glycine buffer, pH 9, 25 mmoles; 25 $\mu$g DNA (salmon sperm); and 195 $\mu$g of fraction C protein. Final volume, 0.20 ml. Incubated 20 minutes at 37°C.

This reduced priming capacity of native DNA was investigated by Bollum (1959a,b) who showed that even at early stages in the purification the calf thymus polymerase uses denatured DNA much more efficiently than native DNA as primer. This is seen clearly in Fig. 17 where several heated and unheated DNA preparations are compared. Except in the case of salmon sperm DNA, the unheated or native DNA preparations are strikingly deficient as primers as compared to the heated preparations. As mentioned previously, Kornberg (1960) has pointed out that some preparations of *E. coli* polymerase prefer denatured DNA preparations as primers. Bollum has suggested that the reason a denatured primer requirement is more easily demonstrated with calf thymus polymerase than with the corresponding enzyme from *E. coli* is that thymus glands contain much less nuclease than *E. coli* and thus polymerase preparations prepared from this tissue have relatively less nuclease activities earlier in the purification procedure. As mentioned before (Section IV,E) the presence of nuclease in the enzyme preparations could convert non-priming (native) DNA to priming (denatured) DNA thus obscuring a requirement for denatured DNA preparations. The implication in all these studies is that the natural primer for DNA synthesis is single-stranded DNA. As in the *E. coli* system, the single-stranded DNA from bacteriophage φX174 is an excellent primer for calf thymus polymerase and requires no prior heating to manifest its activity. Other means of denaturing DNA such as acid or alkali treatment are also effective in converting non-priming to priming DNA (Table XXVI).

TABLE XXVI
COMPARISON OF PRIMER-PRODUCT BASE RATIOS WITH CALF THYMUS POLYMERASE[a]

| | | Treatment | | | |
|---|---|---|---|---|---|
| | Primer DNA | Time (min) | (Temp °C) | Primer ratio | Product ratio |
| | | Experiment 1 | | | |
| A/C | T1 | 2 | 99 | — | 1.3 |
| | T2 | 2 | 99 | 1.9 | 1.8 |
| | T4 | 2 | 99 | 1.9 | 1.9 |
| | T6 | 2 | 99 | 1.9 | 1.8 |
| | Calf thymus | 2 | 99 | 1.3 | 1.8 |
| | *Diplococcus pneumoniae* | 10 | 99 | 1.6 | 1.6 |
| | | Experiment 2 | | | |
| T/C | Calf thymus | 2 | 99 | 1.37 | 1.38 |
| | | 15 | 68 | | Non-priming |
| | | 13 | 78 | | 1.30 |
| | | 15 | 98 | | 1.37 |
| | | 30 | 0, $OH^-$ | | 1.39 |
| | | 30 | 0, $H^+$ | | 1.36 |

[a] From Bollum (1959b).

Calf thymus and *E. coli* polymerase differ in that no "unprimed" reactions have yet been demonstrated with the mammalian enzyme. Calf thymus polymerase does not catalyze the unprimed synthesis of the d-AT or d-GC polymers. These reactions were difficult to discern at first, even with the highly purified *E. coli* enzyme, until the optimal conditions were worked out. It is possible that further attempts to synthesize these products with more purified enzyme preparations from calf thymus under suitable conditions will resolve this difference.

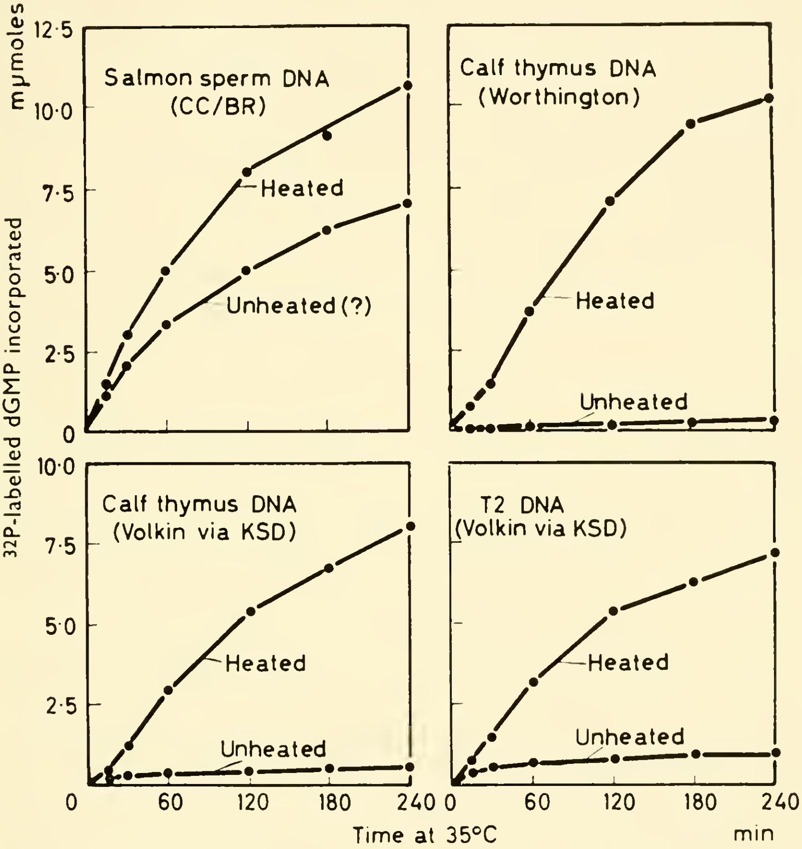


FIG. 17. Comparison of heated and unheated DNA preparations with respect to priming capacity. The question mark indicates the uncertainty concerning the manufacturer's statement that "certain less gentle modifications" are used in the preparative procedure. (From Bollum, 1959a.)

A second difference between the two enzymes, more difficult to resolve, also involves primer. So far, no chemically synthesized polymer has served as primer for *E. coli* polymerase, and DNase digests of primer DNA are totally inactive in this capacity. Calf thymus polymerase, on the other hand, can use DNase digests as well as synthetic polynucleotides as primers for deoxyribonucleotide incorporation (Bollum, 1960b,c). Polyribonucleotides cannot serve as primers, nor can sonicated DNA unless it is heated after sonication to melt out interchain hydrogen bonds.

Preliminary studies with a series of deoxythymidylate homopolymers of general structure $(pT)_n$ where $n \geqq 3$ showed that single deoxyribonucleotides (incubated as triphosphates) added to the ends of these synthetic oligonucleotides. All of the four triphosphates were effective but dGMP and dTMP tended to form multiple products. Since the incubation mixtures contained relatively large excesses of primer to deoxyribonucleoside triphosphate, the most likely event was the addition of single deoxynucleotides to the ends of the chains. The added deoxyribonucleotides were removed from the polymer by snake venom phosphodiesterase suggesting that authentic phosphodiester linkages were established between the oligonucleotide and the incorporated deoxyribonucleotide. Further investigations using chemically synthesized polythymidylates as primer (Bollum, 1962) have extended this earlier work and have shown that oligonucleotides with greater than two nucleotide residues are required. The 3′-hydroxyl group of the terminal deoxynucleoside residue must be free for the condensation since polymers containing an acetylated 3′-hydroxyl group are inactive as primers. Since single deoxyribonucleotides add to the ends of chains in this reaction, the mechanism seems similar to the "limited" reaction catalyzed by *E. coli* polymerases (Section IV,C). The difference is that the latter enzyme requires polymerized DNA and the deoxynucleotides added seem to conform to the base-pairing relationships of the primer.

These observations are part of the general investigation undertaken by Bollum into what is the minimum length of primer required for replicative or hydrogen bond directed synthesis where base-pairing relationships are obligatory. Preliminary observations by Furlong and Bollum (cited in Bollum, 1962) suggest that replicative synthesis begins when chain length exceeds 20 monomer units.

Bollum has suggested (personal communication) that the addition of specific deoxyribonucleotides to the ends of chemically prepared oligonucleotides might provide a useful means of synthesizing specific sequences for testing in other biological systems.

## VI. Synthesis of DNA in Bacteriophage-Infected *Escherichia coli*

### A. BACTERIOPHAGE-SPECIFIC POLYMERASE

The dramatic appearance of several new enzymatic activities soon after infection of *E. coli* with the T-even bacteriophages has recently been reviewed by Cohen (1961). Many of these new activities explain several observations concerning the DNA metabolism of the infected cell. For example, the induction of a specific enzyme which catalyzes the hydroxymethylation of deoxycytidylate (Flaks and Cohen, 1957, 1959a) accounts for the presence of this unique pyrimidine base in bacteriophage DNA. Also, the marked increase in activity of a dCTPase (Zimmerman and Kornberg, 1961; Koerner *et al.,* 1960) may account for the apparent decrease in dCMP kinase activity in extracts of bacteriophage-infected cells. Kornberg *et al.* (1959), in a study of the effect of bacteriophage infection on DNA synthesis, observed a 10-fold stimulation of polymerase activity in extracts of T2-infected cells but noticed that the observed increase was dependent on the type of primer used for the assay. That is, the increased activity after infection could only be demonstrated if heated DNA was used as primer. Recently, Aposhian and Kornberg (1961, 1962) have conducted a careful study of the polymerase in extracts of infected cells and provide conclusive evidence that the increased activity after infection is due to the formation of a new polymerase.

Polymerase from T2-infected cells was purified 600-fold and was distinguishable from *E. coli* polymerase in the following respects:

1. Levels of antiserum against *E. coli* polymerase causing 100% inhibition have no effect on T2 polymerase and vice versa.
2. *p*-Chloromercuribenzoate ($1.7 \times 10^{-4}\, M$) inhibits T2 polymerase completely whereas *E. coli* polymerase is inhibited 27%.
3. T2 polymerase and *E. coli* polymerase are easily separable by chromatography on a column of phosphocellulose.
4. The two polymerases have distinctly different primer requirements.

Except for this last property the differences between the two enzymes should have no consequences for the *in vitro* synthesis of DNA. Thus, the T2 enzyme is similar to the *E. coli* enzyme in that it requires all four deoxyribonucleoside triphosphates, Mg ions, and DNA. It can also incorporate base analogs into DNA in the same manner as the *E. coli* enzyme provided the analogs are supplied as the deoxyribonucleoside triphosphates.

The most striking property of the T2 enzyme is its characteristic dependence on heated DNA first observed in crude extracts of infected cells. The *E. coli* enzyme and T2 polymerase are compared in Table XXVII; it may be seen that heated DNA is a better primer for the T2 polymerase-catalyzed reaction but is actually somewhat worse for the *E. coli* polymerase reaction. That this difference in behavior is not due to some unknown factor in either the purified T2 or *E. coli* polymerases is shown by the mixture of the two preparations. In this case the values equal the sums of the individual incubations.

TABLE XXVII

Primer Differences of T2 and *Escherichia coli* Polymerases[a]

| Primer[b] | Polymerase ($C^{14}$-deoxynucleotide incorporated) T2 (mμmoles) | *E. coli* (mμmoles) | T2 + *E. coli* (mμmoles) |
|---|---|---|---|
| Expt. 1 | | | |
| Unheated salmon sperm DNA | 0.175 | 0.995 | 1.30 |
| Heated salmon sperm DNA | 0.635 | 0.492 | 1.04 |
| Expt. 2 | | | |
| Unheated salmon sperm DNA | 0.249 | 2.00 | 2.11 |
| Heated salmon sperm DNA | 1.30 | 1.20 | 2.84 |

[a] From Aposhian and Kornberg (1962).

[b] In Experiment 1, 0.09 μg of T2 polymerase or 0.10 μg of *E. coli* polymerase was used. In Experiment 2, twice the amount of enzyme was used. The labeled substrate was $C^{14}$-dGTP.

How these properties of the induced polymerase are significant in the infectious process is not known, but it would not be surprising to learn that these differences are essential in allowing the phage to synthesize DNA in an alien environment.

### B. GLUCOSYLATION OF 5-HYDROXYMETHYLCYTOSINE (HMC)

The composition of the DNA in T-even bacteriophages is unique in that it contains 5-hydroxymethylcytosine instead of cytosine (Wyatt and Cohen, 1953), and it also contains glucose (Jesaitis, 1954; Volkin, 1954; Sinsheimer, 1954). Flaks and Cohen demonstrated that a new enzyme, deoxycytidylate hydroxymethylase (which accounted for the synthesis of this viral pyrimidine; see Section II,C) was induced in the bacteriophage-infected system. Likewise, Kornberg *et al.* (1959, 1961; Zimmerman *et al.*, 1962) have indicated that new enzymes capable of glucosylating HMC-containing DNA were formed soon after infection The elegant experiments of Lehman and Pratt (1960) using a highly purified phosphodiesterase which could degrade the normally DNase-resistant bacteriophage DNA to mononucleotides (Lehman, 1960) delineated the structures and distribution of the glucosylated HMC-nucleotides isolated from T2, T4, and T6 DNA preparations. A compilation of their data is shown in Table XXVIII. If we assume that a specific enzyme is required for each type of glucosylated nucleotide, six enzymes (two for each phage) should be demonstrable. So far, five different enzymes have been purified and characterized (Zimmerman *et al.*, 1962;

Josse and Kornberg, 1962). The only activity not yet found is the enzyme which forms diglycosyl-HMC units in T2 DNA.

UDP-glucose (Caputto *et al.*, 1950) has been shown to be the coenzyme required for the transfer of glucose to the HMC residues in each of the reactions tested (Kornberg *et al.*, 1959). The enzymes will not glucosylate the free deoxyribonucleotide or deoxyribonucleoside triphosphate, but require HMC residues in intact DNA. Conversely, synthetic glucosyl hydroxymethyldeoxycytidine triphosphate in the presence of the other three triphosphates is much less effective than the non-glucosylated triphosphate in supporting DNA synthesis in extracts of infected cells (Koerner *et al.*, 1960). The generalized reaction for the $\alpha$-glucosyl and $\beta$-monoglycosyl transferase may be formulated:

TABLE XXVIII

ANALYSIS OF HMC[a]-RESIDUES IN BACTERIOPHAGE DNA[b]

| Glucose linkages | T2 (%) | T4 (%) | T6 (%) |
|---|---|---|---|
| Non-glucosylated | 25 | 0 | 25 |
| $\alpha$-Glucosyl | 70 | 70 | 3 |
| $\beta$-Glucosyl | 0 | 30 | 0 |
| Diglucosyl | 5 | 0 | 72 |

[a] HMC = hydroxymethylcytosine.

[b] Data from Lehman and Pratt (1960).

$$\text{UDP-glucose} + \text{HMC-DNA} \rightarrow \text{UDP} + \text{glucosyl-HMC-DNA}$$

The HMC-DNA used as acceptor for the glucosyl units is prepared by incubating *E. coli* polymerase with 5-hydroxymethyldeoxycytidine triphosphate, the other three deoxyribonucleoside triphosphates, Mg ions, and primer DNA. Several different types of primer, including DNA from *E. coli*, bacteriophage, and calf thymus have been used with equal success. Alternatively, naturally occurring HMC DNA may be used as acceptor provided the heterologous transferase is being studied. For example, T2 DNA (containing 25% unglucosylated HMC residues) is a good substrate for T4 or T6 $\alpha$-transferase but will not accept glucose using the purified T2 transferase. As a general rule, no transferase will catalyze the further glucosylation of its homologous DNA.

The T2, T4, and T6 $\alpha$-glucosyl transferases are three very closely related proteins which have many properties in common. They are different in certain respects as is evident from an inspection of Table XXIX. The T4 $\beta$-transferase is readily distinguishable from any of the $\alpha$-transferases.

TABLE XXIX

COMPARISON OF THE MONOGLUCOSYL TRANSFERASE[a]

| Property | Transferase | | | |
|---|---|---|---|---|
| | T2$\alpha$ | T4$\alpha$ | T6$\alpha$ | T4$\beta$ |
| Requirement for sulfhydryl | + | + | + | 0 |
| Inhibition by $Mg^{++}$ | + | + | + | 0 |
| Inhibition by tris | 0 | 0 | 0 | + |
| $K_m$: DNA | $3 \times 10^{-5}\ M$ | $2.5 \times 10^{-5}\ M$ | $3 \times 10^{-5}\ M$ | $3.3 \times 10^{-5}\ M$ |
| $K_m$: UDP-glucose | $3.6 \times 10^{-5}\ M$ | $0.7 \times 10^{-5}\ M$ | $1.5 \times 10^{-5}\ M$ | $2.1 \times 10^{-5}\ M$[b]<br>$0.6 \times 10^{-5}\ M$[c] |
| Behavior in phosphate buffer: | | | | |
| $\frac{\text{Activity at pH 6.5}}{\text{Activity at pH 7.0}}$ | 4.0 | 4.0 | 1.0 | 0.8 |
| Glucosylation of T2 DNA | 0 | + | + | + |
| Glucosylation of T4 DNA | 0 | 0 | 0 | 0 |
| Glucosylation of T6 DNA | 0 | 0 | 0 | 0 |

[a] From Josse and Kornberg (1962).

[b] In phosphate buffer.

[c] In Ammediol buffer (2-amino, 2-methyl, 1,3-propanediol).

A study of the reversibility of the reaction with the T2 $\alpha$-transferase (Zimmerman *et al.*, 1962) indicates that glucosyl residues may be removed from glucosylated HMC DNA in the presence of UDP but the equilibrium strongly favors the forward reaction:

$$K = \frac{(\text{glucosyl-HMC})(\text{UDP})}{(\text{HMC})(\text{UDP-glucose})} = 200\text{--}2000$$

Thus, the extent of glucosylation is probably not limited by equilibrium considerations since 25% of the available sites in T2 DNA are unglucosylated, and incubation in the presence of T2 $\alpha$-transferase and UDP-glucose does not further glucosylate it.

What then does control the number and positioning of the glucosyl residues in the different DNA types? The simplest hypothesis is that specific base sequences determine the addition of glucosyl units. Recent experiments by Josse and Kornberg (1962) do not tend to support this notion although they do not exclude it. First of all, as had already been observed by Kornberg *et al.* (1961), T4 $\beta$-glucosyl transferase can glucosylate 100% of the available HMC residues in DNA. We know that T4 DNA contains 70% α- and 30% β-glucosyl linkages. Second, we would expect that natural T4 DNA which has had some α-glucosyl residues removed with the α-transferase and UDP (running the reaction in reverse) should accept only α-glucosyl units, and similarly β-deglucosylated DNA should accept only β-glucosyl units. The data in Table XXX, however, show that no matter which of the enzymes is used to remove the glucose units, both enzymes can put the glucose back in. This would indicate that the sequence around the glucosyl units does not direct the synthesis. Third, a polymer containing only dGMP and HMC nucleotides was synthesized using d-GC polymer as primer. This new polymer presumably had the structure $\text{d-G}\overline{\text{HMC}}$ with all the guanine residues in one strand and all the HMC residues in the other. When this polymer was used as an acceptor for glucosylation, all four of the specific monoglucosyl transferases were active. This would mean that all the enzymes were capable of glucosylating an HMC-residue bounded on both sides by another HMC residue.

TABLE XXX
ADDITION OF GLUCOSE TO "DE-GLUCOSYLATED" T4 DNA[a]

| Transferase used to remove glucose | Transferase used to restore glucose | Extent of addition |
|---|---|---|
| None | T4α | <0.001 |
| | T4β | <0.001 |
| T4α | T4α | 0.024 |
| | T4β | 0.024 |
| T4β | T4α | 0.021 |
| | T4β | 0.040 |

[a] From Josse and Kornberg (1962).

If base sequence is not the determinant for insertion of glucose residues, where else may we look? Josse and Kornberg (1962) point out that the secondary or tertiary structure of the macromolecule may be pertinent here. They indicate the heat-denatured DNA is markedly altered in respect to rate of glucosylation and in some instances extent of glucosylation. Indeed, the decrease in rate of glucosylation is so well correlated with the "melting" of DNA that Zimmerman *et al.* (1962) suggest the possibility of using the rate of glucosylation as a rapid, sensitive index for measuring heat denaturation of DNA.

At present, the mechanism which directs the positioning of glucosyl units throughout the DNA molecule is unknown. Equally mysterious is the function of glucose in bacteriophage DNA.

## VII. CONCLUSIONS

The pathway of DNA synthesis in *E. coli* and calf thymus studied *in vitro* are similar in several respects. We may consider the formation of deoxyribonucleotides as the initial reaction in the metabolic sequence leading specifically to DNA. The reduction of ribonucleotides to the corresponding deoxyribosyl derivatives, the phosphorylation of these derivatives and their subsequent polymerization into the macromolecule are being studied as individual reactions in purified systems.

The biological aspects of DNA synthesis, such as cellular growth and proliferation, genetic continuity, and metabolic control, are dependent on the individual enzymatic steps in this pathway and are therefore influenced by the physical properties of the enzymes and the chemical properties of the reactions. Lord Kelvin said, "When you can measure that of which you speak, and express it in numbers, you know something about it." The detailed biochemical investigations of the individual steps in this sequence are providing the "numbers" indispensable in arriving at a satisfactory description of the biological process and are establishing a foundation for the formulation of hypotheses necessary to this area of research.

## ACKNOWLEDGMENTS

I should like to thank Dr. Arthur Kornberg, Dr. John Josse, and Dr. Fred J. Bollum for letting me use material which had not yet been published.

Thanks are also due to Mrs. Mary E. Backer for expert assistance in the preparation of the manuscript, and to the people in my laboratory for living with me during this period.

This manuscript was written during the tenure of a grant (C-4088) from the U. S. Public Health Service.

## REFERENCES

Adler, J., Lehman, I. R., Bessman, M. J., Simms, E. S., and Kornberg, A. (1958). *Proc. Natl. Acad. Sci. U.S.* **44**, 641.

Aposhian, H. V., and Kornberg, A. (1961). *Federation Proc.* **20**, 361.

Aposhian, H. V., and Kornberg, A. (1962). *J. Biol. Chem.* **237**, 519.

Baldwin, R. L., Inman, R. B., and Wake, R. G. (1961). *Federation Proc.* **20**, 354.

Barner, H. D., and Cohen, S. S. (1954). *J. Bacteriol.* **68**, 80.

Barner, H. D., and Cohen, S. S. (1959). *J. Biol. Chem.* **234**, 2987.

Bello, L. J., Van Bibber, M. J., and Bessman, M. J. (1961a). *J. Biol. Chem.* **236**, 1467.

Bello, L. J, Van Bibber, M. J., and Bessman, M. J. (1961b). *Biochim. et Biophys. Acta* **53**, 194.

Berg, P., and Joklik, W. K. (1953). *Nature* **172**, 1008.

Bertani, E., Haggmark, A., and Reichard, P. (1961). *J. Biol. Chem.* **236**, PC 67.

Bessman, M. J. (1959). *J. Biol. Chem.* **234**, 2735.

Bessman, M. J., and Bello, L. J. (1961). *J. Biol. Chem.* **236**, PC 72.

Bessman, M. J., and Van Bibber, M. J. (1959). *Biochem. Biophys. Research Communs.* 1, 101.

Bessman, M. J., Lehman, I. R., Simms, E. S., and Kornberg, A. (1957). *Federation Proc.* 16, 153.

Bessman, M. J., Lehman, I. R., Simms, E. S., and Kornberg, A. (1958a). *J. Biol. Chem.* 233, 171.

Bessman, M. J., Lehman, I. R., Adler, J., Zimmerman, S. B., Simms, E. S., and Kornberg, A. (1958b). *Proc. Natl. Acad. Sci. U.S.* 44, 633.

Bianchi, P. A., Butler, J. A. V., Crathorn, A. R., and Shooter, K. V. (1961). *Biochim. et Biophys. Acta* 53, 123.

Birnie, G. D., and Crosbie, G. W. (1958). *Biochem. J.* 69, 1P.

Blakely, R. L. (1957). *Biochim. et Biophys. Acta* 24, 224.

Bojarski, R. B., and Hiatt, H. H. (1960). *Nature* 188, 1112.

Bollum, F. J. (1958). *J. Am. Chem. Soc.* 80, 1766.

Bollum, F. J. (1959a). *J. Biol. Chem.* 234, 2733.

Bollum, F. J. (1959b). *In* "The Cell Nucleus" (J. S. Mitchell, ed.), p. 60. Academic Press, New York.

Bollum, F. J. (1960a). *J. Biol. Chem.* 235, 2399.

Bollum, F. J. (1960b). *J. Biol. Chem.* 235, PC 19.

Bollum, F. J. (1960c). *Federation Proc.* 19, 305.

Bollum, F. J. (1962). *J. Biol. Chem.*, in press.

Bollum, F. J., and Potter, V. R. (1958). *J. Biol. Chem.* 233, 478.

Bollum, F. J., and Potter, V. R. (1959). *Cancer Research* 19, 561.

Bollum, F. J., Anderegg, J. W., McElya, A. B., and Potter, V. R. (1960). *Cancer Research* 20, 138.

Bowman, H. G. (1957). *Nature* 180, 1182.

Canellakis, E. S. (1959). *Proc. 4th Intern. Congr. Biochem., Vienna, 1958*, p. 75.

Canellakis, E. S., and Måntsavinos, R. (1958). *Biochim. et Biophys. Acta* 27, 643.

Canellakis, E. S., Jaffe, J. J., Mantsavinos, R., and Krakow, J. S. (1959). *J. Biol. Chem.* 234, 2096.

Caputto, R., Leloir, L. F., Cardini, C. E., and Paladini, A. C. (1950). *J. Biol. Chem.* 184, 333.

Cohen, S. S. (1961). *Federation Proc.* 20, 641.

Cohen, S. S., and Weed, L. L. (1954). *J. Biol. Chem.* 209, 789.

Cohen, S. S., Flaks, J. G., Barner, H. D., Loeb, M. R., and Lichtenstein, J. (1958). *Proc. Natl. Acad. Sci. U.S.* 44, 1004.

Colowick, S. P., and Kalckar, H. M. (1943). *J. Biol. Chem.* 148, 117.

Crosbie, G. (1960). *In* "The Nucleic Acids" (E. Chargaff and J. N. Davidson, eds.), Vol. III, pp. 337–342. Academic Press, New York.

Cunningham, L., Catlin, B. W., and de Garilhe, M. P. (1956). *J. Am. Chem. Soc.* 78, 4642.

Davidson, E. A. (1959). *Biochim. et Biophys. Acta* 33, 238.

de Garilhe, M. P., and Laskowski, M. (1956). *J. Biol. Chem.* 223, 661.

Dische, Z. (1955). *In* "The Nucleic Acids" (E. Chargaff and J. N. Davidson, eds.), Vol. I, p. 285. Academic Press, New York.

Doty, P. (1957). *J. Cellular Comp. Physiol.* 49, 27.

Dunn, D. B., and Smith, J. D. (1954). *Nature* 174, 305.

Edmunds, M. (1958). *Federation Proc.* 17, 215.

Elwyn, D., and Sprinson, D. B. (1954). *J. Biol. Chem.* 207, 467.

Flaks, J. G., and Cohen, S. S. (1957). *Biochim. et Biophys. Acta* 25, 667.

Flaks, J. G., and Cohen, S. S. (1959a). *J. Biol. Chem.* 234, 1501.

Flaks, J. G., and Cohen, S. S. (1959b). *J. Biol. Chem.* 234, 2891.

Flaks, J. G., Lichtenstein, J., and Cohen, S. S. (1959). *J. Biol. Chem.* 234, 1507.

Freese, E. (1959). *J. Mol. Biol.* 1, 87.

Friedkin, M. (1959a). *Federation Proc.* 18, 230.

Friedkin, M. (1959b). *In* "The Kinetics of Cellular Proliferation" (F. Stohlman, Jr., ed.), p. 97. Grune & Stratton, New York.

Friedkin, M., and Kalckar, H. M. (1950). *J. Biol. Chem.* 184, 437.

Friedkin, M., and Kornberg, A. (1957). *In* "The Chemical Basis of Heredity" (W. D. McElroy and B. Glass, eds.), pp. 609–613. Johns Hopkins Press, Baltimore, Maryland.

Friedkin, M., and Roberts, D. (1954). *J. Biol. Chem.* 207, 245, 257.

Gray, E. D., Weissman, S. M., Richards, J., Bell, D., Keir, H. M., Smellie, R. M. S., and Davidson, J. N. (1960). *Biochim. et Biophys. Acta* 45, 111.

Green, M., and Cohen, S. S. (1957). *J. Biol. Chem.* 225, 387.

Greenberg, G. R., and Somerville, R. L. (1962a). *Proc. Natl. Acad. Sci. U.S.* 48, 247.

Greenberg, G. R., Somerville, R. L., and De Wolf, S. (1962b). *Proc. Natl. Acad. Sci. U.S.* 48, 242.

Hammarsten, E., Reichard, P., and Saluste, E. (1950). *J. Biol. Chem.* 183, 105.

Hecht, L. I., Potter, V. R., and Herbert, E. (1954). *Biochim. et Biophys. Acta* 15, 134.

Heidelberger, C., Chandhuri, N. K., Danneberg, P., Mooren, D., Griesbach, L., Duschinsky, R., Schuitzer, R. J., Plever, E., and Scheimer, J. (1957). *Nature* 179, 663.

Heppel, L. A., Markam, R., and Hilmoe, R. (1953). *Nature* 171, 1151.

Hiatt, H. H., and Bojarski, T. B. (1960). *Biochem. Biophys. Research Communs.* 2, 35.

Hilmoe, R. J. (1960). *J. Biol. Chem.* 235, 2117.

Humphreys, G. K., and Greenberg, D. M. (1958). *Arch. Biochem. Biophys.* 78, 275.

Hurwitz, J. (1959). *J. Biol. Chem.* 234, 2351.

Jesaitis, M. A. (1954). *Microbial Genetics Bull.* 10, 16.

Josse, J., and Kornberg, A. (1962). *J. Biol. Chem.* 237, 1968.

Josse, J., Kaiser, A. D., and Kornberg, A. (1961). *J. Biol. Chem.* 236, 864.

Kalle, G. P., and Gots, J. S. (1961). *Biochim. et Biophys. Acta* 51, 130.

Keck, K., Mahler, H. R., and Fraser, D. (1960). *Arch. Biochem. Biophys.* 86, 85.

Khorana, H. G., Tener, G. M., Razell, W. E., and Markham, R. (1958). *Federation Proc.* 17, 312.

Kielley, R. K. (1961). *Biochim. et Biophys. Acta* 53, 150.

Kirkland, R. J., and Turner, J. F. (1959). *Biochem. J.* 72, 716.

Klenow, H., and Anderson, B. (1957). *Biochim. et Biophys. Acta* 23, 94.

Klenow, H., and Lichtler, E. (1957). *Biochim. et Biophys. Acta* 23, 6.

Koerner, J. F., Smith, M. S., and Buchanan, J. M. (1960). *J. Biol. Chem.* 235, 2691.

Kornberg, A. (1957). *In* "The Chemical Basis of Heredity" (W. D. McElroy and B. Glass, eds.), p. 589. Johns Hopkins Press, Baltimore.

Kornberg, A. (1959). *Harvey Lectures* Ser. 50, 83.

Kornberg, A. (1960). *Science* 131, 1503.

Kornberg, A., Lieberman, I., and Simms, E. S. (1955). *J. Biol. Chem.* 215, 389.

Kornberg, A., Zimmerman, S. B., Kornberg, S. R., and Josse, J. (1959). *Proc. Natl. Acad. Sci. U.S.* 45, 772.

Kornberg, S. R., Zimmerman, S. B., and Kornberg, A. (1961). *J. Biol. Chem.* 236, 1487.

Krebs, H. A., and Hems, R. (1953). *Biochim. et Biophys. Acta* **12**, 172.

Kunitz, M. (1950). *J. Gen. Physiol.* **33**, 349.

Laskowski, M., Hagerty, G., Laurila, U. (1957). *Nature* **180**, 1182.

Lehman, I. R. (1959). *Ann. N.Y. Acad. Sci.* **81**, 745.

Lehman, I. R. (1960). *J. Biol. Chem.* **235**, 1479.

Lehman, I. R., and Pratt, E. A. (1960). *J. Biol. Chem.* **235**, 3254.

Lehman, I. R., Bessman, M. J., Simms, E. S., and Kornberg, A. (1958a). *J. Biol. Chem.* **233**, 163.

Lehman, I. R., Zimmerman, S. B., Adler, J., Bessman, M. J., Simms, E. S., and Kornberg, A. (1958b). *Proc. Natl. Acad. Sci. U.S.* **44**, 1191.

McNutt, W. S. (1958a). *J. Biol. Chem.* **233**, 189.

McNutt, W. S. (1958b). *J. Biol. Chem.* **233**, 193.

MacNutt, W. S. (1952). *Biochem. J.* **50**, 384.

Maley, G. F., and Maley, F. (1959). *J. Biol. Chem.* **234**, 2975.

Maley, F., and Ochoa, S. (1958). *J. Biol. Chem.* **233**, 1538.

Mandel, L. R., and Borek, E. (1961). *Biochem. Biophys. Research Communs.* **6**, 138.

Manson, L. A., and Lampen, J. O. (1951). *J. Biol. Chem.* **191**, 95.

Markham, R., and Smith, J. D. (1953). *Biochem. J.* **52**, 552.

Marmur, J., and Doty, P. (1959). *Nature* **183**, 1427.

Michelson, A. M. (1958). *Nature* **182**, 1502.

Moore, E. C., and Hurlbert, R. B. (1960). *Biochim. et Biophys. Acta* **40**, 371.

Ochoa, S., and Heppel, L. A. (1957). *In* "The Chemical Basis of Heredity" (W. D. McElroy and B. Glass, eds.), pp. 633–637. Johns Hopkins Press, Baltimore.

Phear, E. A., and Greenberg, D. M. (1957). *J. Am. Chem. Soc.* **79**, 3737.

Potter, R. L., Schlesinger, S., Buettner-Janusch, V., and Thompson, L. (1957). *J. Biol. Chem.* **226**, 381.

Pouyet, J. (1952). *Compt. rend. acad. sci.* **234**, 152.

Racker, E. (1952). *J. Biol. Chem.* **196**, 347.

Reichard, P. (1957). *Acta. Chem. Scand.* **11**, 11.

Reichard, P. (1958). *Biochim. et Biophys. Acta* **27**, 434.

Reichard, P. (1959). *J. Biol. Chem.* **234**, 1244.

Reichard, P. (1960). *Biochim. et Biophys. Acta* **41**, 368.

Reichard, P., and Rutberg, L. (1960). *Biochim. et Biophys. Acta* **37**, 554.

Reichard, P., Baldesten, A., and Rutberg, L. (1961). *J. Biol. Chem.* **236**, 1150.

Remy, C. N., Remy, W. T., and Buchanan, J. M. (1955). *J. Biol. Chem.* **217**, 885.

Roll, P. M., Weinfeld, H., and Carrol, E. (1956). *J. Biol. Chem.* **220**, 455.

Rose, I. A., and Schweigert, B. S. (1953). *J. Biol. Chem.* **202**, 635.

Sable, H. Z., Wilber, P. B., Cohen, A., and Kane, M. R. (1954). *Biochim. et Biophys. Acta* **13**, 156.

Scarano, E. (1958). *Biochim. et Biophys. Acta* **29**, 459.

Scarano, E., and Tolarico, M. (1959). *Boll. soc. ital. biol. sper.* **35**, 97. Cited in Maley and Maley (1959).

Schachman, H. K., Lehman, I. R., Bessman, M. J., Adler, J., Simms, E. S., and Kornberg, A. (1958). *Federation Proc.* **17**, 304.

Schachman, H. K., Adler, J., Radding, C. M., Lehman, I. R., and Kornberg, A. (1960). *J. Biol. Chem.* **235**, 3242.

Shapiro, S., and Chargaff, E. (1960). *Nature* **188**, 62.

Shugar, D. (1960). *In* "The Nucleic Acids" (E. Chargaff and J. N. Davidson, eds.). Vol. III, p. 39. Academic Press, New York.

Siekevitz, P., and Potter, V. R. (1953). *J. Biol. Chem.* **200**, 187.

Singer, M. F., Hilmoe, R. J., and Heppel, L. A. (1958). *Federation Proc.* **17**, 312.

Sinsheimer, R. L. (1954). *J. Biol. Chem.* **208**, 445.

Sinsheimer, R. L. (1954). *Science* **120**, 551.

Sinsheimer, R. L. (1959). *J. Mol. Biol.* **1**, 43.

Smellie, R. M. S. (1960). *Biochem. J.* **77**, 168.

Somerville, R., Ebisuzaki, K., and Greenberg, G. R. (1959). *Proc. Natl. Acad. Sci. U.S.* **45**, 1240.

Sueoka, N. (1961). *J. Mol. Biol.* **3**, 31.

Swartz, M. N., Trautner, T. A., and Kornberg, A. (1962). *J. Biol. Chem.* **237**, 1961.

Trautner, T. A., Swartz, M. N., and Kornberg, A. (1962). *Proc. Natl. Acad. Sci. U.S.* **48**, 449.

Volkin, E. (1954). *J. Am. Chem. Soc.* **76**, 5892.

Wahba, A. J., and Friedkin, M. (1961). *J. Biol. Chem.* **236**, PC 11.

Warner, R. C. (1957). *J. Biol. Chem.* **229**, 711.

Watson, J. D., and Crick, F. H. C. (1953a). *Nature* **171**, 737.

Watson, J. D., and Crick, F. H. C., (1953b). *Cold Spring Harbor Symposia Quant. Biol.* **18**, 23.

Weissman, J. M., Smellie, R. M. S., and Paul, J. (1960). *Biochim. et Biophys. Acta* **45**, 101.

Wyatt, G. R., and Cohen, S. S. (1953). *Biochem. J.* **55**, 774.

Zimmerman, S. B., and Kornberg, A. (1961). *J. Biol. Chem.* **236**, 1480.

Zimmerman, S. B., Kornberg, S. R., and Kornberg, A. (1962). *J. Biol. Chem.* **237**, 512.

*Chapter* II

# The Replication and Organization of DNA in Chromosomes

J. HERBERT TAYLOR

*Department of Botany and*
*Department of Zoology,*
*Columbia University,*
*New York, New York*

I. Introduction . . . 65
II. A Morphological Picture of the Chromosome . . . 67
  A. Mitotic Chromosomes . . . 67
  B. Giant Chromosomes . . . 68
III. Replication of DNA . . . 72
  A. Chromosomal Level . . . 72
  B. Molecular Level . . . 77
IV. Exchanges between DNA Subunits and Genetic Recombination . . . 82
  A. Chromosomal Level . . . 82
  B. Molecular Level . . . 86
V. Sequences of Replication of DNA in Chromosomes . . . 88
VI. Chromosome Reproduction in Relation to DNA Replication and Protein Synthesis . . . 89
  A. General Comments . . . 89
  B. Types of Chromosomal Aberrations Produced at Various Stages of the Cell Cycle . . . 90
  C. Healing and Reunion of Broken Chromosomes . . . 92
VII. Molecular Models for Organization of DNA into Chromosomes . . . 96
  References . . . 108

## I. INTRODUCTION

Reproduction is a unique characteristic of living systems. Obviously it can occur at many levels of organization. The largest hiatus in our understanding of reproduction involves problems of the establishment and maintenance of organization above the ordinary molecular level. Most of the enzymatic reactions of the cell can be demonstrated in cell-free systems and at least one virus (tobacco mosaic) can be dissociated and reassembled. However, we are far from being able to reconstruct the simplest cell, much less a multicellular organism. In principle the problem is not beyond the realm of possibility for we have the concept that all of the complexity is traceable to properties of the constituent molecules and their disposition in space and time. Therefore, we attack the problem by trying to understand the organization of the cell's constituent molecular systems.

In molecular genetics we might begin by trying to understand the reproduction and organization of the genetic structures themselves—the chromosomes. The term is used here in its broadest sense to designate the entities in which is recorded the major part of the genetic information of cells and the more complex viruses. The smallest of these appear to consist of a single molecule (Chapter III). Whether we refer to a large chromosome as a molecule or not may be a matter of semantics, but evidence indicates that it has some of the properties usually associated with molecules—a definite size, atomic integrity over long periods of time, and a regular sequence of its constituent parts which are bonded together by rather stable chemical bonds.

A basic hypothesis of molecular genetics is that the synthesis of DNA is a process of replication, i.e., that new molecules are made by copying a complementary sequence of the various nucleotides from an original template which is not destroyed in the process. Much of the evidence supporting this concept has been presented in Chapter I. The demonstration of the atomic integrity of two DNA subunits in a chromosome (Taylor *et al.*, 1957) and later in the various particles of DNA isolated from chromosomes (Meselson and Stahl, 1958) was a significant contribution to this concept, but these discoveries posed a larger question—a problem of organization, packaging, and sorting that may not be solved soon. The conserved subunits in the DNA particles are very likely the individual polynucleotide chains, although this interpretation has been challenged (Cavalieri and Rosenberg, 1961a). Likewise the two conserved subunits of a chromatid are almost certainly the same as those in its fragments (the isolated DNA molecules or particles). Some of the questions that are unresolved concern the continuity of the DNA along the chromosome, the nature of linkers if these exist, the singleness or degree of multiplicity of the chromonema, and the pattern of packaging, folding, or coiling in the various structures that a chromosome assumes during the phases of the cell cycle. In addition, the mechanisms which regulate function and replication almost certainly have a basis in the structure and organization of a chromosome, but the details are largely unknown.

Three possible arrangements include most of the conceptual pictures which have been presented: (*1*) a multistranded complex of DNA, protein, and perhaps other constituents with many DNA double helices forming the axis of the chromonema; (*2*) some regular arrangement of molecules of DNA or of nucleoprotein particles linked together in tandem with non-DNA linkers, or alternately attached by their ends to a shorter central axis; and (3) one long continuous DNA double helix with its associated protein folded and coiled to form the visible chromonema of a typical chromosome.

First a morphological picture of a chromosome will be presented briefly. Then the evidence on replication at the chromosomal and molecular levels will be summarized. After examining the information available on the properties of the subunits, their exchanges and reunion during genetic recombination, and their repair following radiation-induced breakage, the results will be interpreted in terms of a molecular model of a chromosome.

## II. A MORPHOLOGICAL PICTURE OF THE CHROMOSOME

### A. MITOTIC CHROMOSOMES

Chromosomes present a changing pattern in relation to the cell division cycle, and during interphase when some of the most important and significant changes are occurring they present little structure which has yielded to analysis with either the light or the electron microscope. Perhaps it will be appropriate to begin a consideration of the chromosome at anaphase when it is composed of a single chromatid. In some favorable preparations these rather flexible rods can be seen to consist of helically coiled chromonemata. Usually they appear to have a single chromonema, but in some fixed preparations, probably due to treatment with acid fixatives, the chromatids appear double. In these there may be two helically coiled chromonemata (half-chromatids) usually very poorly separated as if the coils were interlocked and perhaps the axes otherwise bonded together. Fixation in hot water may also reveal the doubleness, but *in vivo* a doubleness has never been demonstrated. In the living state anaphase chromatids appear to be solid cylinders. An exception may occur in certain coccids which have chromosomes with a multiple or diffuse spindle attachment (Hughes-Schrader, 1940) and in certain symbiotic protozoa of termites (Cleveland, 1949). In the secondary spermatocytes of the coccids the half-chromatids may be completely separated. In the symbiotic protozoa mentioned above half-chromatids are reported to be separated in some cells and not in others of the same species. Whether these move apart during fixation or actually separate in the living cell is not clear. In addition, it is not clear that the half-chromatids in these cases are comparable to the half-chromatids of other organisms. However, it may be noted that the agents (fixatives) which supposedly produce this partial separation would break hydrogen bonds as well as denature proteins, but would not be expected to break covalent bonds.

At the end of anaphase there may be a further condensation of the chromosomes but following this phase they expand somewhat in telophase and the coiled chromonemata become less tightly coiled. The chromosomes then appear to fuse and nearly fill the reformed interphase nucleus. As the nuclei grow, and particularly in cells that are not to divide again, the chromosomes may not continue to fill the nucleus. Their disposition in most interphase nuclei is difficult to follow either by the use of the light microscope or the electron microscope.

In dividing cells there is typically a gap ($G_1$) of a few hours before DNA replication begins (see Chapter IV for further details). Synthesis ($S$ period) then occurs over a period of several hours followed by another gap ($G_2$) before the next division.

By the earliest prophase stages at which chromosomes can be distinguished, they can be seen to consist of two chromatids which continue to shorten during prophase. The chromonemata can be seen to be helically coiled in some types of cells. The gyres increase in diameter and their number apparently decreases during prophase as the chromatids shorten.

The coils usually show best in meiotic chromosomes. In addition, the chromomeres of the classical cytologists, which have sometimes been equated with poorly resolved coils, are most striking at the prophase of the first meiotic division. Although chromomere patterns along the chromonemata vary in cells at different stages of development or in different cells in a species (Lima-de-Faria *et al.*, 1959), the mapping of pachytene and diplotene chromosomes indicates that chromomeres are significant structural features comparable to the bands along the giant chromosomes of Diptera. The ultimate shape and dimensions of chromonemata, in other words, the packing pattern of the nucleoprotein in the chromosomes, is still largely unknown. Investigation is difficult because the structural details are generally below the level of resolution of the light microscope and often above the size level at which the thin sections required for resolution in electron microscopy can give a complete picture. An additional difficulty appears to be an instability of any tertiary structure in the nucleoprotein components of chromosomes.

## B. GIANT CHROMOSOMES

From a study of structures such as prophase and anaphase chromosomes one would easily reach the conclusion that there are many strands along their axes at the molecular level. A chromonema with a diameter of $0.1–0.3\,\mu$ is several orders of magnitude larger than the extended nucleohistone fiber of $30\,\text{Å}$. However, in a few special cases chromosomes elongate far beyond the length observed at early meiotic prophase (leptotene). The giant chromosomes of the Diptera provide one example, but their structure is very complex and they contain many chromonemata. In addition, since these giant chromosomes are end products of differentiation, there may be some reason to question their exact homology with mitotic chromosomes. However, in the oöcyte of amphibians a very much elongated chromosome exists which does revert to a regular mitotic type at the following division (Fig. 1A). These chromosomes are 500–800 $\mu$ long in diplotene (Gall, 1956), but change into chromosomes only 15–20 $\mu$ in length at later stages. In addition to this shortening by a factor of 40–50 along the axis there are numerous loops which are retracted into the main body of the chromosome. When extended these would give the chromosome a contour length of several centimeters. A few years ago Callan and McGregor (1958) showed that DNase (deoxyribonuclease) would rapidly sever the loops of these isolated lampbrush chromosomes. Although some cytologists had long maintained that the loops were part of the chromatid axis, this was the first clear demonstration that the genetic material, DNA, passed through and maintained the continuity of the loops. In addition to the DNA of the loops, the chromomeres along the axis, which consist of two closely associated chromatids, were known to be Feulgen positive and therefore to contain much of the DNA. This elongation to several hundred times the length of a metaphase chromosome in itself leads one to suspect that much of the bulk of even the early meiotic prophase chromosomes is due to some type of folding or coiling of the much longer axis of one or a few DNA strands. The difficulty of determining the structure of the loops lies in the fact that a great amount of RNA and protein is attached to them. In spite of this neither RNase (ribonuclease) nor proteases will sever the loops. Recently Miller (1962) has been able to dissolve off this protein and RNA coating with concentrated KCl solution and to reveal the fine fibril, presumably DNA, which remains. Electron micrographs reveal its diameter to be sometimes as small as 40–50 Å (Fig. 1B). Gall (1962) has been able to obtain evidence, by the rate at which the loops are broken by DNase, that the structural axis is composed of a single DNA double helix. Since the loops are part of the axis of single chromatids, these elongated structures which are among the largest chromosomes in the animal kingdom are shown to have a single DNA double helix as their linear component. DNase also produces breaks between the chromomeres along the axis of the paired chromatids with the kinetics predicted for a structure held together by a pair of DNA double helices. If all of these results are correctly interpreted we finally have the demonstration that the polymer maintaining the linear

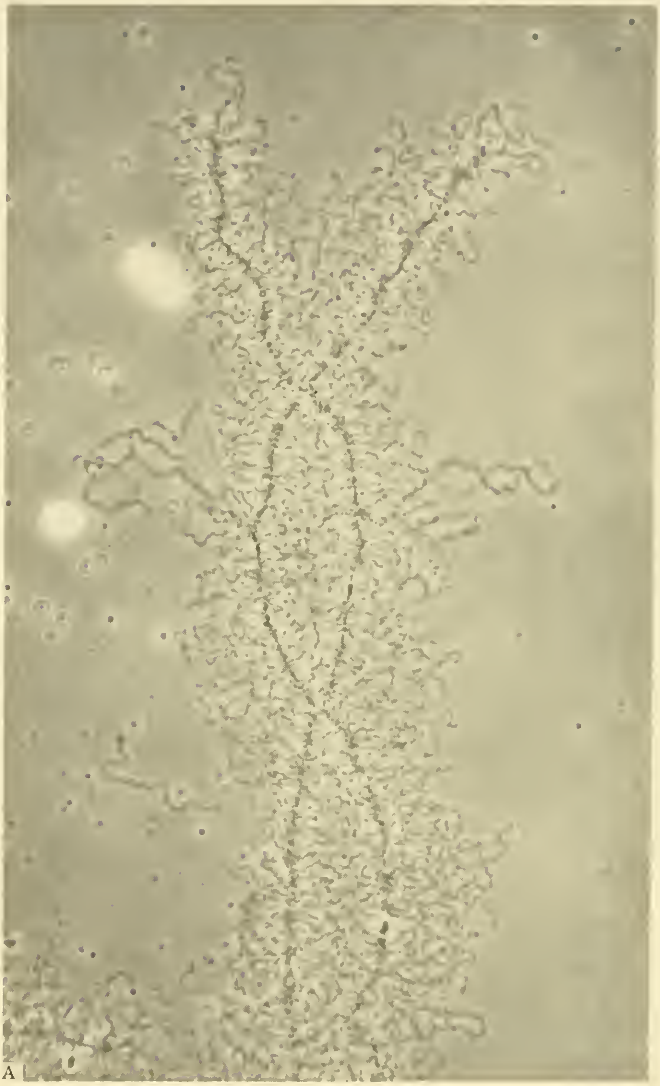


FIG. 1A. Phase photograph of unfixed lampbrush bivalent chromosome from an oöcyte of the newt, *Triturus viridescens* (× 650). Present evidence indicates that DNA is the axial element which maintains the continuity of the material in each of the loops extending from the main axis. (Original photograph contributed by J. G. Gall.)

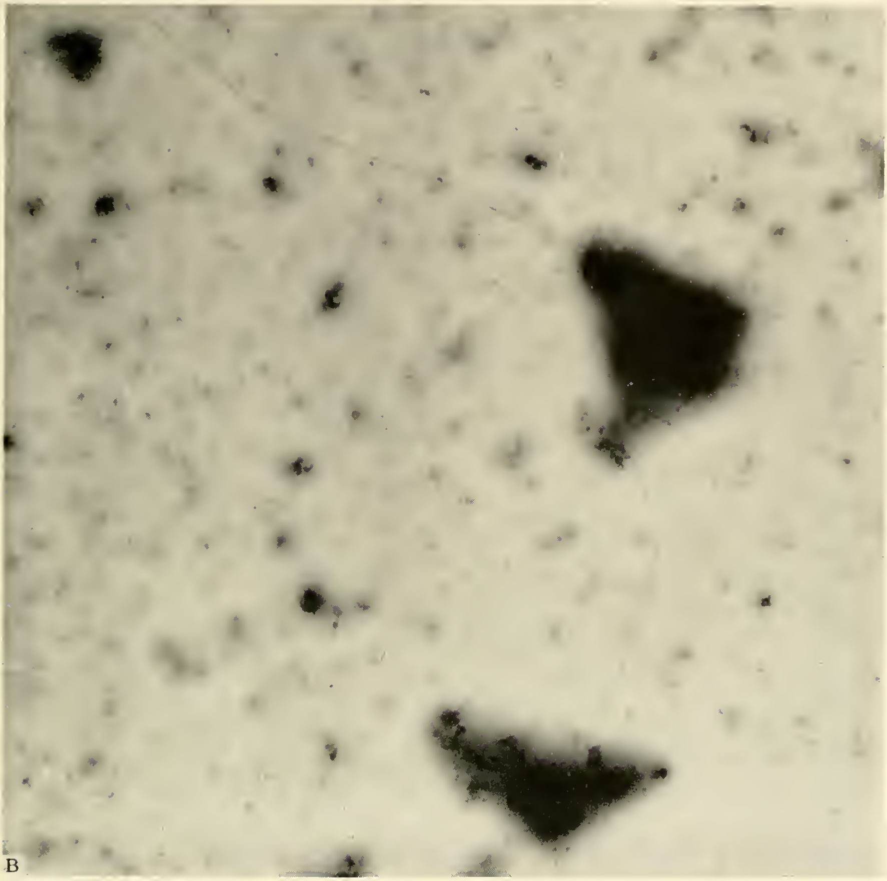


FIG. 1B. Electron micrograph of a small portion of a loop of a lampbrush chromosome. Most of the protein and RNA has been dissolved away with a saturated solution of KCl after placing an isolated chromosome on Formvar film. The dissolution was interrupted by adding saturated $(NH_4)_2SO_4$ at pH 5.1 followed by staining with uranyl acetate (× 75,000). Thinner portion of axial fiber is 40–50 Å in diameter. (Original photograph contributed by O. L. Miller.)

integrity of a chromatid is, at some stage anyway, DNA, and that a single pair of polynucleotide chains from this axis. Linkers along the axis are not ruled out but neither are any revealed by these experiments. However, long stretches of peptides or RNA which might serve as linking material appear unlikely from the fact that proteases and RNase do not break the loops.

In addition to the much extended chromosomes of the amphibian oöcytes, other oöcytes and spermatocytes of many species have chromosomes of a less extreme lampbrush type. These chromosomes as well as those of the amphibian oöcyte were called lampbrush chromosomes because some early cytologists who studied them thought the extensions from the axis were not loops but bristles like those of a brush. Actually they may be bristles instead of loops at the earlier stages, during pachytene, for example (Taylor, 1962). This point has not been established. However, Moses (1956) demonstrated that an electron-dense axial element exists in the meiotic chromosomes and that when fixed in neutral osmium tetroxide the lateral extensions were fine filaments 50–100 Å in diameter which are often paired where they radiate from the axis. That they contain DNA could be demonstrated by staining adjacent thick sections of the same chromosome by the Feulgen reaction. Moses (1960) proposed a model which is consistent with this structure as well as that observed in the nuclei of amphibian oöcytes described above.

The axial element cannot be demonstrated after diplotene by which time the bristles, if that is what they are rather than loops, are being folded or otherwise packaged in the condensing chromosome. Further details of the chromosome's organization is not discernible with the electron microscope as it approaches the metaphase state.

## III. REPLICATION OF DNA

### A. CHROMOSOMAL LEVEL

Until Watson and Crick (1953a) proposed a specific model for DNA, ideas concerning its mechanism for transmission of genetic information were rather vague. However, following the demonstration by Avery *et al.* (1944) that the transforming principle of *Pneumococcus* was DNA, the idea that DNA was the genetic material began to influence the design of experiments. The finding of a constancy in amount of DNA per genome (Boivin *et al.*, 1948; Mirsky and Ris, 1949) and finally the experiments by Hershey and Chase (1952) which indicated that only the DNA of phage entered the bacterial cell made the case more convincing. Watson and Crick (1953a) took the information available on bond distances and bond angles, on the most likely tautomeric forms for the bases, on the X-ray diffraction pictures which suggested a helix, and the chemical analyses, which indicated that DNA contained a 1:1 ratio of adenine to thymine and of guanine to cytosine, and constructed a model in which the purine and pyrimidine bases were directed toward the axis of the helix. By allowing two hydrogen bonds between each of these base pairs in opposing antiparallel chains which rotated to form a double helix with one gyre per 10 nucleotide pairs, a stable rather stiff molecular model was obtained with the phosphoric acid residues at the periphery of the structure.

The complementarity of the two chains and the specificity of the hydrogen bonds suggested a mechanism for precise replication (Watson and Crick, 1953b). If the chains came apart, each contained the hydrogen bonding sites to line up the nucleotides of a complementary chain. This type of replication would yield two identical molecules (Chapter I). With this proposal of a precise model, the foundations were laid for experimental tests. Techniques and suitable biological materials had to be found. By taking the most optimistic view of all the findings, namely, (*1*) that DNA was the genetic material of the cell, (*2*) that all of the DNA of a chromosome was exactly reproduced once each division cycle, (*3*) that the proposed structure was correct for all of the DNA, and (*4*) that the Watson-Crick scheme for its replication was correct, one could reach the conclusion that one chain of each original DNA molecule in a chromosome or in a whole cell would be passed to each daughter nucleus at mitosis. A new complementary chain would likewise be built along each and passed on to the descendants. If a label (radioactive isotope) were present only in the original DNA or available only for the synthesis of the new chains, the distribution could be predicted. The possibilities were formally stated by Delbrück and Stent (1957): (*1*) if the original double helix remained intact and somehow produced a completely new two-stranded molecule, the scheme would be conservative; (*2*) if the chains separated without breakage and each served as a template for the formation of a complementary chain the scheme would be semi-conservative; and (*3*) if the chains were broken or destroyed in the process the scheme would be dispersive.

Attempts to follow the distribution of carbon-14-labeled DNA in whole cells at anaphase by autoradiography might have provided a partial answer. At least a clear result in a cell with a few chromosomes could have distinguished between a conservative scheme versus a semi-conservative or dispersive replication. However, the published results (Plaut and Mazia, 1956) did not provide a clear answer although they were interpreted by the authors to favor a conservative scheme.

To get a definitive answer a technique that could follow the distribution of label at the chromosomal level was required. This was provided for large chromosomes by autoradiography with tritium-labeled thymidine, first prepared and used by Taylor *et al.* (1957). Because of its low energy beta emission tritium gives better resolution for autoradiography than any other isotope. In addition, compared to carbon-14 its shorter half-life (12.3 years) allows one to prepare compounds with a much higher specific activity. Since thymidine, which was labeled by an exchange reaction (Taylor, 1960a), is also a selective label for DNA, other components of the cells would not be labeled. Because the intracellular pool of thymidine and its derivatives is small in most cells, the removal of the label after one cycle of replication would allow one to follow the distribution of labeled DNA through several division cycles and thereby distinguish a semi-conservative replication from the dispersive one.

The results for root cells of *Vicia* were clear (Fig. 2A). After one or a part of a cycle of duplication with thymidine-$H^3$, the two chromosomes descended from an original unlabeled one were both labeled (Taylor *et al.*, 1957). After one more duplication in the absence of the labeled

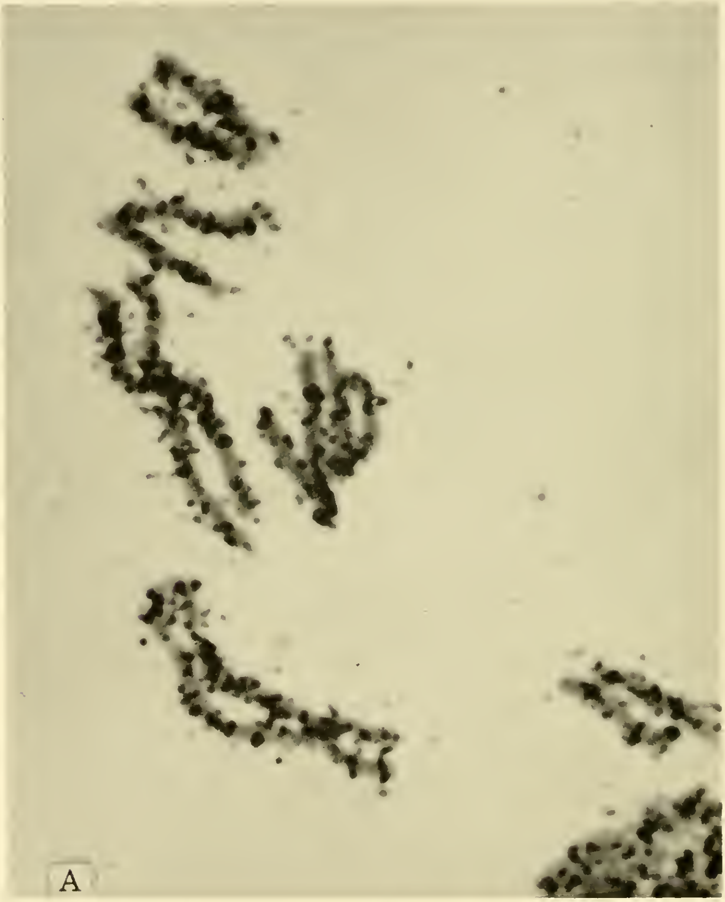
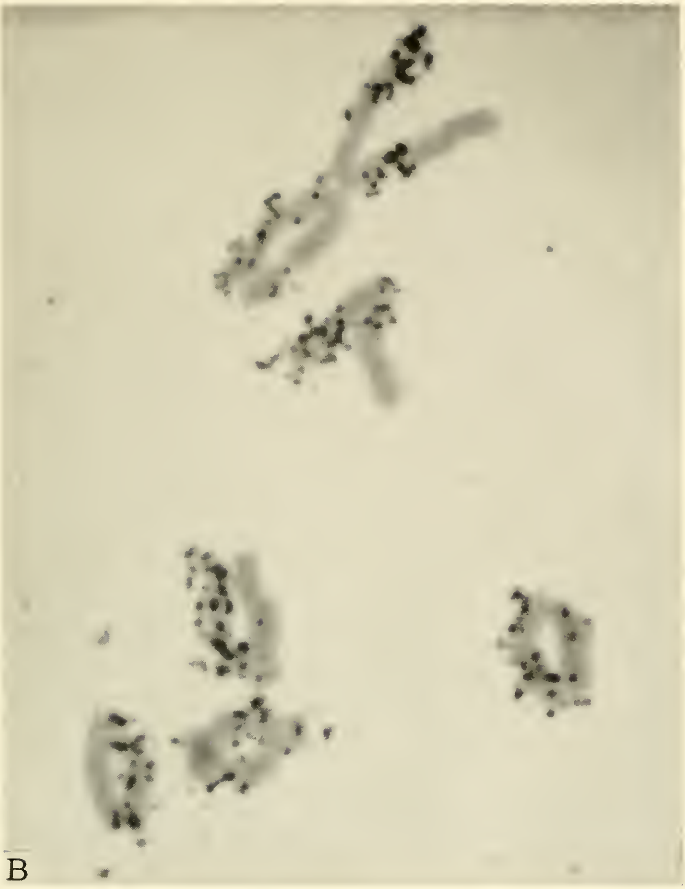


FIG. 2A. Autoradiogram of *Vicia* chromosomes at division after one replication in which thymidine-$H^3$ was incorporated into the DNA. FIG. 2B (see facing page). Autoradiogram of a similar group of chromosomes after one replication with thymidine-$H^3$ and one additional replication in an environment free of labeled precursors of DNA (× 3200).



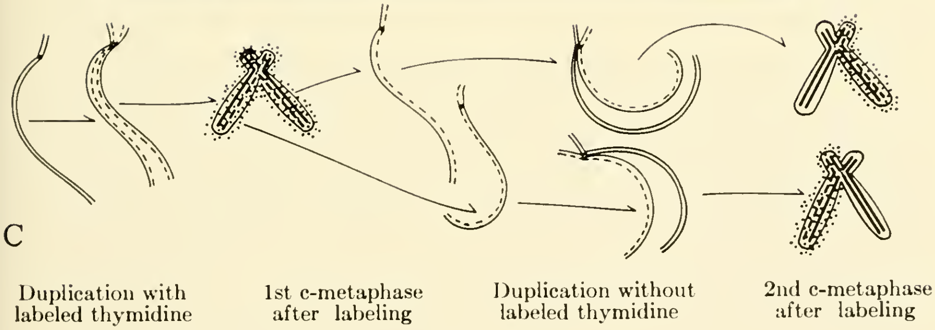

C
Duplication with labeled thymidine
1st c-metaphase after labeling
Duplication without labeled thymidine
2nd c-metaphase after labeling

Fig. 2B. For legend see page 74.

Fig. 2C. Diagrammatic representation of the distribution of subunits of DNA during chromosome duplication. Although the subunits are not resolved by microscopic examination, they are represented and distributed as indicated by the observed pattern of labeling in Figs. 2A and B. Broken lines represent labeled subunits and unbroken lines represent unlabeled subunits. The dots represent grains in the autoradiograms. (From Taylor *et al.,* 1957.)

thymidine, the labeled chromosomes produced one labeled and one unlabeled descendant (Fig. 2B). This is the result predicted for semi-conservative replication of DNA (Fig. 2C). The examination of other species of plants, *Bellevalia* (Taylor, 1958a), *Crepis* (Taylor, 1958b), and *Allium* (Gall, 1958a) confirmed the regular semi-conservative behavior at the chromosomal level. The first experiments were monitored by the use of colchicine. It was used to accumulate cells at division and to identify the cells which had gone through two cycles. By placing roots in colchicine after labeling and leaving them in this solution until labeled tetraploid cells appeared, we could identify chromosomes which had gone through one cycle of duplication after the one in which they were labeled. These cells regularly showed one labeled chromatid paired with one unlabeled chromatid (Fig. 2B). Chromatid exchanges between the labeled and the unlabeled chromatids tended to obscure the picture, but these later proved of value in analyzing the organization of chromatids (see below).

A report by LaCour and Pelc (1958) criticized the use of colchicine and reported that it had an effect on the segregation of the labeled DNA in *Vicia* chromosomes. However, a re-examination of the problem by Woods and Schairer (1959) failed to show any effect of colchicine. In addition the studies on *Bellevalia* and *Crepis* (Taylor, 1958a,b) were carried out with and without colchicine. The results with respect to segregation of labeled DNA were the same in both instances. Rare instances of apparent equal segregation of labeled DNA after the second replication in colchicine-treated cells were reported by Taylor (1958a). These could result from exchanges among the four chromatids which were the descendants of a pair of labeled chromatids. Colchicine, which probably keeps such pairs together, would be expected to increase the frequency of this event. In addition an unequal labeling of sister chromatids may appear at the first division in the first cells to reach division. Usually these have relatively less radioactivity than the majority of labeled division figures. With a low number of grains the significance of any apparent asymmetry of labeling in sister chromatids is difficult to assess. Recent studies (Taylor, unpublished) with thymidine-$H^3$ of higher specific activity than was available in early experiments reveal what appears to be a significant asymmetry in the labeling of sister chromatids. This is interpreted as evidence for the out-of-phase replication of the two subunits of a chromatid.

Examination of two mammalian species (human and Chinese hamster) and an insect (grasshopper) has revealed the semi-conservative distribution of tritium-labeled DNA (Taylor, 1960b and unpublished).

Forro and Wertheimer (1960) have also carried out extensive studies on the segregation of tritium-labeled DNA in whole cells of *Escherichia coli*. Cells of *E. coli* 15T⁻ were exposed to thymidine-$H^3$ of high specific activity for less than one division time and then removed from the label and allowed to divide a number of times. The distribution of labeled DNA was very heterogeneous in clones derived from single young cells. The pattern of segregation indicated that two large DNA-containing units regularly separate at cell division. These units may then be perpetuated intact for many successive cell divisions, but are subject to fragmentation at a finite probability. These observations are consistent with the view that the two subunits represent the single chromosome, which replicates in a semi-conservative fashion, but is subject to an occasional sister chromatid exchange as observed in higher forms.

These results certainly rule out any dispersive mechanism of replication, but they cannot be taken as definitive proof of a semi-conservative scheme at the molecular level. Because of the uncertainties concerning the structure and organization of chromosomes it was conceivable, although highly unlikely, that the two subunits of DNA in a chromosome represented a higher level of organization than that of the molecule of DNA.

### B. MOLECULAR LEVEL

The first evidence that replication might be semi-conservative at the molecular level was provided by Levinthal (1956). In an autoradiographic study of the distribution of phosphorus-32-labeled DNA from phage T2, he obtained evidence that about 40% of the DNA, which remained in one piece when the phage was broken by osmotic shock, was replicated semi-conservatively. The particles obtained from cells infected with $P^{32}$-labeled phage yielded DNA fragments after similar treatment with one-half as much $P^{32}$ as the original phage. Doubts concerning the validity of the results were raised when other methods of analysis failed to reveal these large pieces (see Chapter III; also Davison and Levinthal, 1961). It now appears that the large pieces were probably some sort of artifact and therefore the significance of the original experiments remains unclear. However, more convincing results at the molecular level were soon presented.

Meselson *et al.* (1957) published a technique for separating DNA molecules of different densities in a cesium chloride gradient at high centrifugal force. They soon applied the technique to a study of the distribution of labeled DNA during replication in *Escherichia coli* (Meselson and Stahl, 1958). The cells were grown for a number of generations in $N^{15}$-labeled medium. The heavy DNA of these cells could be banded separately from the regular $N^{14}$-containing DNA (Fig. 3).

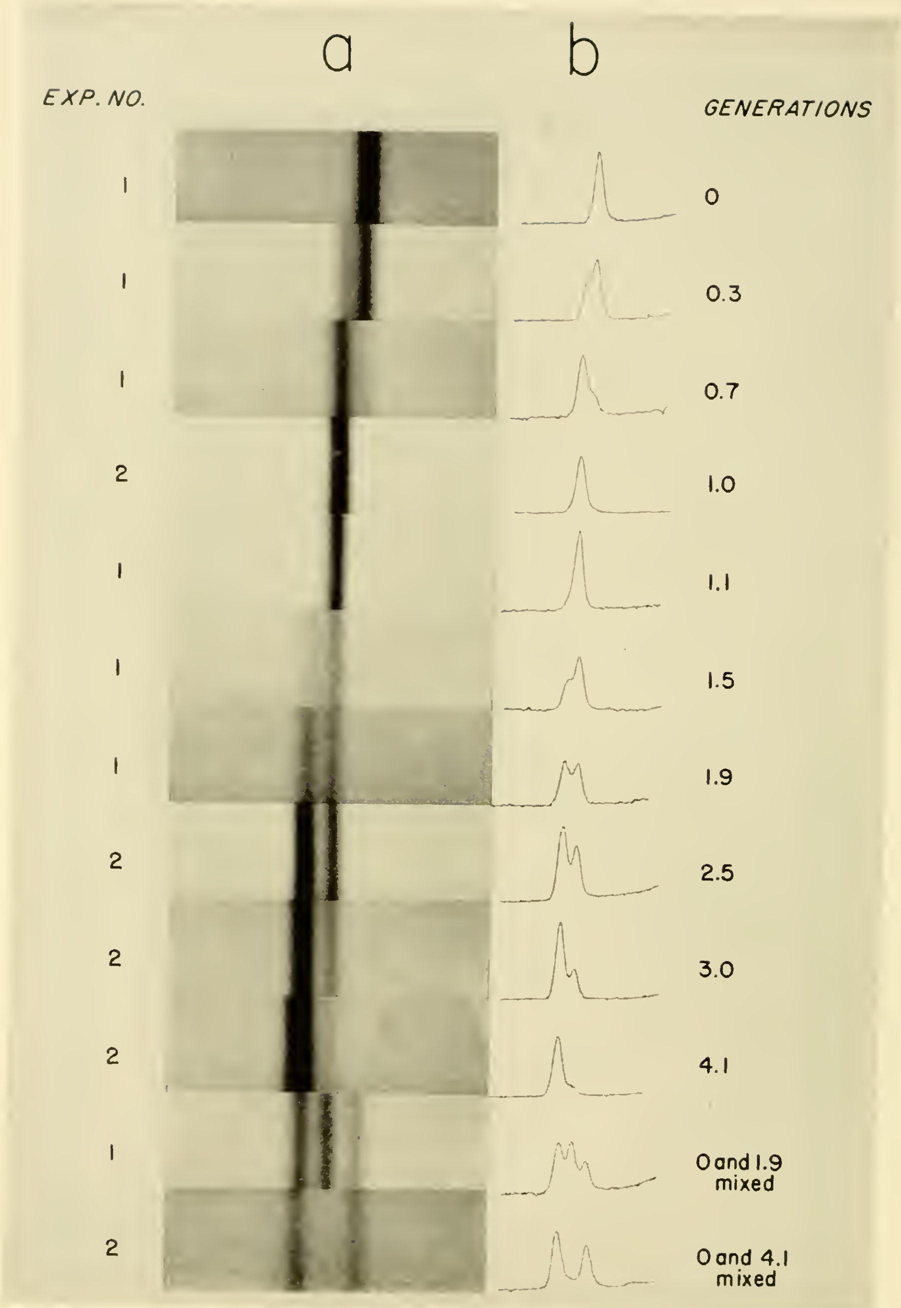


FIG. 3. (*a*) Ultraviolet absorption photographs showing DNA bands resulting from density-gradient centrifugation of lysates of *E. coli* sampled at various times after the addition of an excess of $N^{14}$ substrate to a culture of cells labeled with

They changed a rapidly growing culture of *E. coli* labeled with $N^{15}$ to a medium in which the nitrogen source contained the lighter isotope, $N^{14}$. They took samples at intervals as replication continued. When lysates of these samples of cells were centrifuged in the Spinco Model E and photographed at equilibrium, as shown in Fig. 3, the behavior of the labeled DNA undergoing replication was revealed. A new band of DNA with a density intermediate between the $N^{14}$- and $N^{15}$-labeled material appeared first. By the end of one cycle of replication all of the DNA was in the band with this intermediate density, while the original material with a high density had disappeared. As growth continued a band corresponding in density to $N^{14}$-DNA began to appear. This $N^{14}$-DNA continued to increase in amount, but the band of intermediate density did not disappear. These results indicated very clearly that replication is semi-conservative at the molecular level (Fig. 4). Meselson and Stahl (1958) made another observation of significance. When the DNA of intermediate density was heated for 30 minutes at 100°C and centrifuged again, it formed two separate bands. Therefore, the DNA of intermediate density was composed of hybrid particles in which the separable strands had different densities. The question raised was difficult to answer. Did the replication involve the unwinding of the two polynucleotide chains of a Watson-Crick double helix? Originally this problem of how two chains with one twist every 34 Å could separate was one of major concern. Further consideration and some calculations of the energy requirements and the time that might be required for unwinding have indicated that the problem was more imaginary than real. Levinthal and Crane (1956) showed that the energy requirement and the speed to be expected, if the double helix rotated like a cable, were well within the limits available for cellular processes. More recent experiments and calculations have indicated that no problem exists if the replicating units are similar in size to the isolated particles of DNA (Miller, 1961; Geiduschek, 1961).

Similar results for DNA of mammalian cells were soon provided. Djordjenic and Szybalski (1960) and Simon (1961) used bromouracil as a density label for DNA and obtained evidence for a semi-conservative replication in human cells (HeLa strain) in culture. Evidence was also

---

$N^{15}$. Each photograph was taken after 20 hours of centrifugation at 44,770 rpm in a solution of CsCl; the density of the solution increases to the right. The time of sampling was measured in units of generation time as indicated on the right. (*b*) Microdensitometer tracings of the DNA bands adjacent to each photograph. The density of each band can be compared by reference to the two lowest frames which show the separation of the mixtures indicated on the right. (From Meselson and Stahl, 1958.)

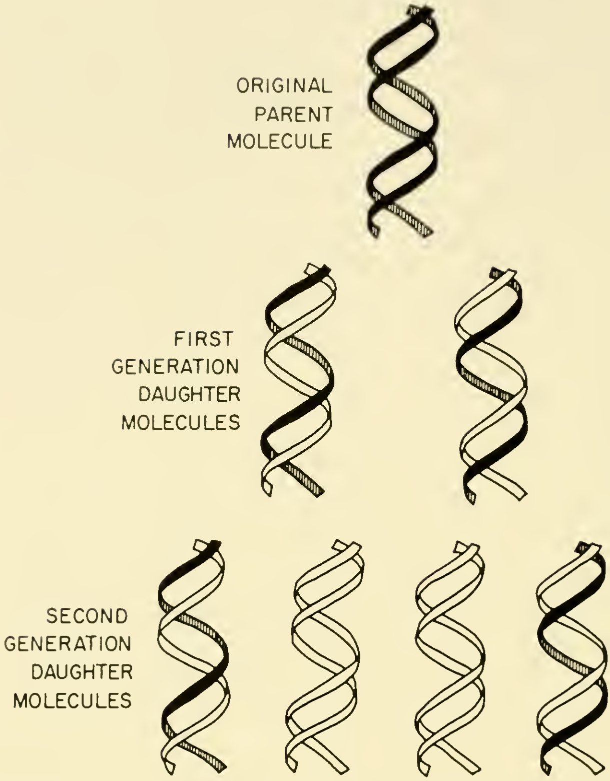


FIG. 4. Illustration of the segregation of DNA chains according to the scheme proposed by Watson and Crick. If the original labeled DNA contains two chains of high density ($N^{15}$), the daughter molecules after one replication in a medium containing $N^{14}$ would have one chain with $N^{15}$ and one with $N^{14}$. After a second replication in the $N^{14}$ medium one-half of the molecules would still have one chain with $N^{15}$ and one with $N^{14}$, but the other half would have two chains with only $N^{14}$. All molecules after one generation in the $N^{15}$ medium should have a density intermediate between the original $N^{15}$ DNA and the $N^{14}$ DNA as shown in the fourth frame of Fig. 3. (From Meselson and Stahl, 1958.)

available that distribution in this same cell line was semi-conservative at the chromosomal level (Taylor, 1960b). More recent studies have shown that the particles of intermediate density from human cells can also be separated into a heavy and light subunit (Chun and Littlefield, 1961).

A semi-conservative replication has also been demonstrated for phage λ (Meselson and Weigle, 1961), phage T7 (Meselson, 1960), and for the alga *Chlymadomonas* (Sueoka, 1960).

The most serious challenge to the simple interpretation of experiments on semi-conservative replication which indicate separation of the two complementary chains of a double helix comes from a series of papers by Cavalieri and Rosenberg (1961a,b,c). These authors maintain, primarily on the basis of the kinetics of molecular weight reduction by enzymes, that DNA extracted from rapidly growing cells is four-stranded. In particular their results indicate that the particles of DNA isolated from *E. coli* such as those used by Meselson and Stahl (1958) are composed of paired double helices held together by hypothetical biunial bonds. According to their interpretation these are the bonds broken by heating and the separation of hybrid DNA observed by Meselson and Stahl involved not the unwinding of the polynucleotide chains, but the separation of a heavy Watson-Crick double helix from a double helix of lower density. Later Hall and Cavalieri (1961) published electron micrographs of *E. coli* DNA which indicated that at least some particles of DNA consisted of paired double helices. In addition, the mass per unit length of the particles, calculated from measurements in the electron micrographs and molecular weights determined by the light-scattering technique, indicated that the particles contained between 3 and 4 polynucleotide chains. On the other hand, there are a number of experiments which clearly indicate that agents known to break hydrogen bonds cause the separation of the strands of DNA particles. Among these are heat, low or high hydrogen ion concentration, and agents that compete for hydrogen-bonding sites or destroy the water layer of hydration, such as urea and formamide (Doty, 1960; Marmur and Doty, 1961; and Schildkraut *et al.*, 1961). The separated strands will recombine if the population of molecules is not too heterogeneous. In addition, hybrid molecules can be produced by heating and slowly cooling DNA's with different isotopic densities.

Whatever may prove to be the significance of these conflicting interpretations of experiments on strandedness of different DNA samples, the Watson-Crick scheme for replication, in which specific pairing of complementary bases is the principal role of the template, is likely to dominate our thinking and the design of experiments in the foreseeable future (see Chapters III, IV, and V). The semi-conservative mode of replication is likewise not a trivial quirk of nature. A phenomenon so universal would not be preserved in evolution so consistently unless it played a fundamental role in reproduction. The only basis on which it makes any sense is the separation of complementary units as a part of the process of replication. One of the most puzzling problems of chromosome reproduction as stated earlier is the packaging and sorting of the tremendously long chromosomal strands if the axial element is a single DNA helix. A model that folds and forms temporary H linkers during replication has been proposed (Taylor, 1962) to provide a sorting and control mechanism. Fragmentation of such a structure could give the multistranded DNA reported for rapidly proliferating cells (Cavalieri and Rosenberg, 1961a,b). Further details of the model will be presented in Section VII.

## IV. Exchanges between DNA Subunits and Genetic Recombination

### A. CHROMOSOMAL LEVEL

Sister chromatid exchanges had been indicated by the behavior of ring chromosomes in maize (McClintock, 1938; Schwartz, 1955) but the autoradiograms of tritium-labeled chromosomes (Fig. 2) provided unmistakable visual evidence for their occurrence in other species (Taylor, 1958a, 1959a). Their frequency and pattern could be determined in well-spread preparations of chromosomes at the second division after labeling Exchanges were not revealed in the first division because all chromatids were labeled. When the exchanges became visible they had to be interpreted in terms of events during two duplication cycles—the one in which labeling occurred and the succeeding one.

An analysis was made from root cells of *Bellevalia,* which has four pairs of large chromosomes (Taylor, 1958a). Attention will be directed to the exchanges in the largest chromosome of the complement. Since we are dealing with diploid cells, there are two of the large homologous chromosomes, each composed of a single chromatid before duplication. Each chromatid contains two subunits of DNA that extend throughout the length—these are the units conserved during replication. After duplication the chromosomes consist of two chromatids, each with two subunits, one new labeled subunit and one unlabeled original. If the subunits are analogous to the polynucleotide chains of the Watson-Crick double helix, they will be unlike. Reciprocal exchanges between the two chromatids will require the union of labeled subunits with unlabeled subunits in each chromatid because only like chains can join (Fig. 5). Since each chromatid will have a labeled subunit along the entire length at the following metaphase, there will be no visible evidence of the exchange. However, when the two subunits separate and replicate at the next duplication (referred to below as the second interphase), one chromatid will have the labeled subunit up to an exchange point and beyond that locus its sister chromatid will contain the labeled subunit. The exchanges will now be revealed by autoradiography at the succeeding metaphase, because silver grains will be over one chromatid from the end to an exchange and over its sister chromatid from that point to another exchange. The metaphase chromosome derived from the other sister chromatid of this originally labeled pair will show an exchange at the same locus (the twin exchanges referred to hereafter). The two metaphase chromosomes descended from the other homolog in these diploid cells will have twin exchanges at different loci if an exchange occurred during its labeling, i.e., the exchanges will be at different loci because they are independent events.

Unpaired (single) exchanges will appear as a result of exchanges during or following the duplication at the second interphase. These will be independent events in all four chromosomes and will be revealed immediately when the chromosomes reach metaphase.

The tendency of exchanges to occur as twins was evident in the first sample of cell examined (Table I). In fact these were more numerous than expected compared to the single, unpaired exchanges. If exchanges at the second interphase occurred at the same frequency as in the first interphase, i.e., the one in which the label was incorporated, the ratio would be one twin pair (product of a single reciprocal exchange) to each two single exchanges produced at the second interphase. Since there are four reproducing chromosomes at second interphase if all descendants of the original two homologs are considered, the chance for single exchanges is increased by a factor of two compared to first interphase.

When the two subunits of a pair of chromatids are unlike as assumed in the prediction of the 1:2 ratio of twin and single exchanges, all of the exchanges in the first interphase will produce twins and all of those in the second interphase will produce singles. On the other hand, if the two subunits were alike as would be true if each were represented by a Watson-Crick double helix, only one-fourth of the exchanges of the first interphase would yield twins and, of course, no twins would be produced from exchanges at the second interphase. Therefore, the predicted ratio would be one twin pair to each 10 singles (Taylor, 1958a).

The results of the first experiment revealed a 2:1 ratio, i.e., more twins than predicted but a ratio not at all compatible with results predicted in a chromosome consisting of identical subunits. The deviation appeared to be the result of an increased exchange frequency at the interphase when labeling occurred or a decreased one at the second interphase. The latter proved to be correct. Colchicine was present at the second interphase, but not during most of the first. It was applied just before the labeled cells, which were being analyzed, reached division. By placing the roots in colchicine at the same time labeling began, i.e., along with the thymidine-$H^3$, the ratio was shifted in the predicted direction, i.e., 1:1 (Table I). In a third experiment the colchicine was added 2 hours before labeling began. In this instance the ratio was

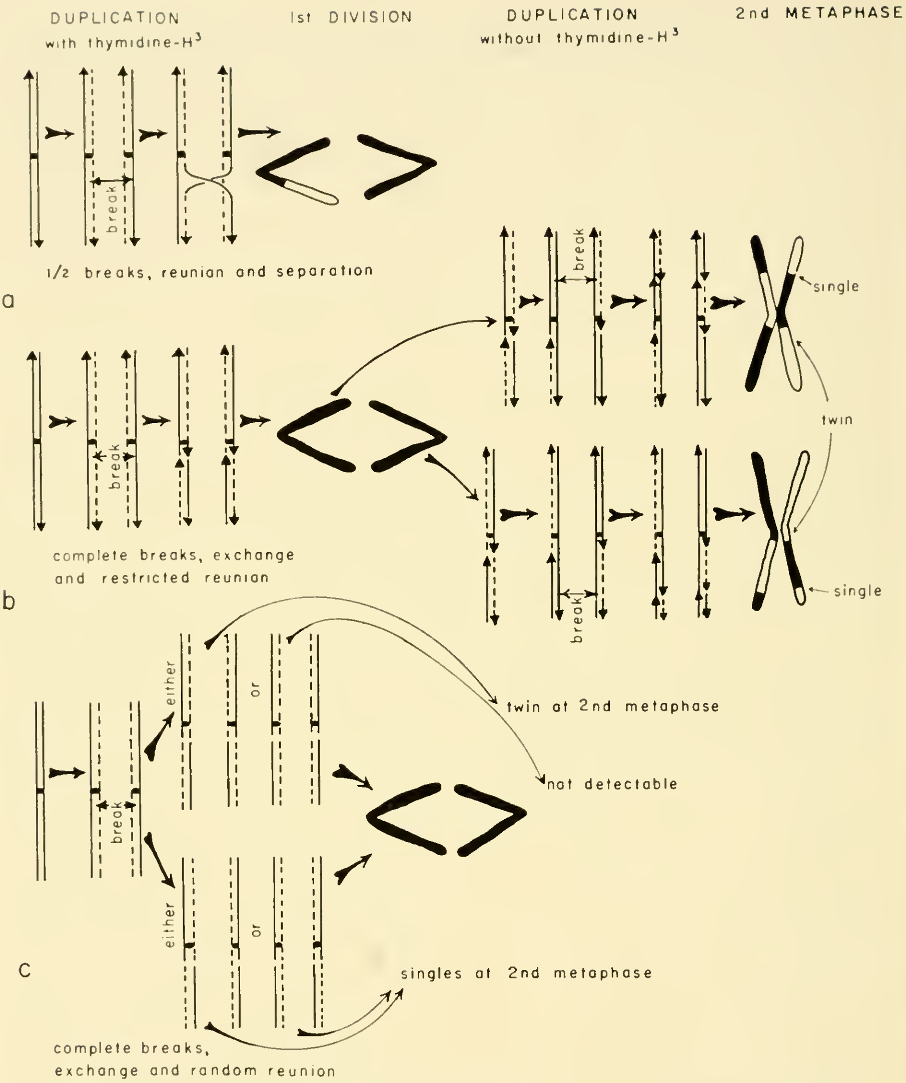

DUPLICATION
with thymidine-$H^3$
1st DIVISION
DUPLICATION
without thymidine-$H^3$
2nd METAPHASE
break
1/2 breaks, reunion and separation
a
break
single
twin
single
break
complete breaks, exchange
and restricted reunion
b
break
either
or
twin at 2nd metaphase
not detectable
break
either
or
c
singles at 2nd metaphase
complete breaks,
exchange and random reunion

FIG. 5. Diagrams showing the predicted results of chromosome reproduction and sister chromatid exchange in which DNA was labeled during one replication with thymidine-$H^3$; the dashed lines represent labeled subunits and the chromosomes shown in solid black represent regions showing grains in autoradiograms. The unbroken lines represent unlabeled subunits and the chromosome regions shown in outline represent regions without grains in the autoradiograms as shown in Fig. 2. (*a*) The predicted results of exchanges involving only one subunit in each chromatid (these were undetected in the experiments and perhaps do not occur). (*b*) The predicted results of exchanges when the two subunits are restricted in their reunion by a difference in sense analogous to the two chains of a DNA double helix.

very nearly the predicted one, and furthermore, the absolute frequency of exchanges decreased in the first interphase as colchicine was added at earlier times. The frequency of exchanges at the second interphase remained nearly constant in all three experiments.

TABLE I
FREQUENCY OF SINGLE AND TWIN EXCHANGES IN THE TETRAPLOID CHROMOSOME SETS OF *Bellevalia romana*[a]

| | | | | Frequency per chromosome | |
|---|---|---|---|---|---|
| Experiment[b] | Number of chromosomes | Twin exchanges | Single exchanges | First interphase | Second interphase |
| 1 | 72 (18 cells) | 36 | 15 | 1.00 | 0.21 |
| 2 | 112 (28 cells) | 39 | 36 | 0.70 | 0.28 |
| 3 | 96 (24 cells) | 14 | 26 | 0.29 | 0.27 |

[a] Data for only the largest chromosome are given.
[b] Data for the first two experiments are from Taylor (1958a,b).

The conclusion reached was that the two subunits of a chromatid are unlike. In addition, the exchanges appear to occur during DNA replication, for colchicine must be applied at that time in order to get the maximum effect on their frequency. While these results strongly suggest a chromosome model in which the subunit is a single polynucleotide chain of the DNA double helix, they do not exclude models in which the axis of the chromosome is a tandemly linked sequence of DNA helices (Taylor, 1959a,b).

Sister chromatid exchanges have been observed in all species analyzed to date. These include several genera of plants, *Vicia, Bellevalia, Crepis, Allium,* two mammalian species, Chinese hamster, and human cells in culture, and an insect, the grasshopper (Taylor, 1958a,b; 1959a; 1960b; and unpublished). None of these except *Bellevalia* have been analyzed for the occurrence of twin exchanges. From what is known of the survival of ring chromosomes in the Diptera, it appears that sister chromatid exchanges may be absent or occur at a very low frequency in this group.

---

(Exchanges are assumed to occur with an equal frequency during the first duplication with thymidine-$H^3$ and during the second duplication without thymidine-$H^3$. All exchanges in the first duplication produce twins and the frequency of single exchanges to twin pairs is 2:1; see text for explanation.) (*c*) The predicted frequency of single and twin exchanges if the two subunits reunite at random. (Since single exchanges could result from events at both the first and second duplication, the frequency of single exchanges would be higher, at least 10:1.) (From Taylor, 1958a.)

However, to date it has not been possible to test this hypothesis directly by means of labeling these rather small chromosomes.

### B. MOLECULAR LEVEL

The exchanges seen at the level of the chromosome would not be detectable at the molecular level by any means now known. For example, if the chromosome were a tandemly linked series of DNA particles with molecular weights of 10–20 × $10^6$, a single large chromosome would contain more than 100,000 of these particles. Therefore, one or two exchanges per chromosome per replication cycle even if all exchanges involved breakage of the DNA would not be detectable at the molecular level. One must conclude that the chromosome is composed of two remarkably stable structural units. This stability may be more apparent than real because the chromosome may have higher orders of structural organization and repair mechanisms to mend breaks (see Section VII). In any case, the DNA of bacteriophages appears to be considerably less stable in linear organization during transmission. Kozinski (1961) has reported a dispersion of labeled phage T4 DNA that would indicate much more frequent breakage than appears to occur in large chromosomes (see Chapter III for further details). The two subunits of phage T7 appear to be much more stable in their linear organization. Most of the subunits remain intact through several replications and Meselson (1960) could demonstrate that the distribution of DNA was semi-conservative. Phage λ also undergoes exchange rather infrequently compared to phage T4. Replication has been shown to be semi-conservative, but exchanges occur often enough to be detectable by density-labeling experiments (Meselson and Weigle, 1961; Kellenberger *et al.*, 1961).

Meselson and Weigle (1961) used the heavy isotopes of nitrogen and carbon to label one parent. In crosses of these with phages containing the light isotopes and appropriate genetic markers, exchange could be shown even between phage chromosomes which did not undergo replication. At low multiplicity of infection (0.1 phage/cell) the original DNA was distributed to progeny phage particles as predicted if the original "phage chromosome" was composed of two subunits of DNA and at the first replication one heavy subunit was contributed to each new chromosome. These phage chromosomes of hybrid density were not dispersed by exchange during the subsequent replications enough to obscure the essential semi-conservative distribution. However, an actual breakage and exchange even of non-replicating "phage chromosomes" could be demonstrated at a high multiplicity of infection (5–7 phage/cell) with appropriate genetic markers. Phage were also recovered which contained chromosomes with only one of the original labeled subunits. By the use of the genetic markers phage could also be recovered with the densities expected if they contained a single heavy subunit that had undergone an exchange between the two markers. The markers used were *mi* (at one end of the genetic map) and *c* which is 5 units away. The remainder of the map is 12.5 units. The density of both the conserved particles (containing unreplicated chromosomes) and the semiconserved particles with a crossover between the markers was that predicted if the chromosome lost an average of about 14% of its heavy label, i.e., breaks occurred at random between the loci *c* and *mi*.

Kellenberger *et al.* (1961) found mutants of phage λ having different densities. The density differences were apparently due to insertion of extra DNA at the *b2* locus in one mutant and at the *b5c* locus in the other. The markers are 7.5 units apart, i.e., about 43% of the map. The two mutant phages could be separated by centrifugation in a cesium chloride gradient. Recombination between the two markers yielded particles denser than either parent and also lighter than either parent. When they produced the light phage, labeled its DNA with $P^{32}$, and crossed it with the heavy phage, most of the $P^{32}$ transmitted remained with the light parental particles. However, the particles recombinant for the density markers contained considerably more $P^{32}$ than non-recombinant heavy particles. They interpreted the results to show clearly a breakage and exchange of DNA associated with recombination. They also interpreted the evidence to indicate that most of the $P^{32}$ transferred to the particles non-recombinant for the density markers was received by breakage and exchange distal to the two markers.

These experiments make a significant contribution to our understanding of the mechanism of recombination in phage chromosomes and perhaps in larger chromosomes as well. The experiments are comparable to the classic experiments performed by Stern (1931) and Creighton and McClintock (1931) which demonstrated that genetic recombination in higher organisms involved an exchange of chromosome segments. They go further than the chromosome experiments in that they demonstrate that copy choice in replication is not the only mechanism of recombination. In fact there is still no direct evidence that copy choice occurs. However, it is not ruled out by the experiments and is still a convenient mechanism to explain aberrant segregation, perhaps as a part of breakage and exchange events. By putting together the evidence available on physical exchange between chromatids (sister chromatid exchange made visible in tritium-labeled chromosomes) and these recently reported data on phage chromosomes, we can formulate the hypothesis that a variable amount of exchange characteristic for each type of genetic organization occurs in those forms with the smallest chromosomes as well as those with the largest ones.

## V. Sequences of Replication of DNA in Chromosomes

As already noted, DNA replication is a discontinuous process during the cell cycle. In cells of higher forms the period of synthesis may be 6 to 8 hours or perhaps longer in some cells. Even in cells where replication is completed in much less time, the period is long compared to the time required to synthesize most large molecules. Therefore, the question of the sequence of replication and its control has often been raised. Evidence at the molecular level is very scarce. The mechanism in which nucleotides are added at the 3′ OH ends by DNA polymerase (Chapter I) would appear to require a sequential addition. In addition, the failure to find particles with the whole range of densities when a density label is introduced into a replicating system has led to the suggestion that there are few growing points at any moment during the *S* period. However, the only direct evidence relating to the sequence of replication has been obtained at the level of the chromosome by use of tritium-labeled DNA. These studies now clearly indicate that there are control mechanisms that operate to control the replication among chromosomes within a single nucleus as well as the sequence or order of replication within a single chromosome. This type of control probably has a structural basis (see Section VII on chromosome models).

The first indication of asynchronous replication within chromosomes was the observation that the regions around the centromeres were replicated late in root cells of *Crepis* (Taylor, 1958b). Later Lima-de-Faria (1959) found that the X-chromosome of a grasshopper continued its replication longer than the autosomes. Taylor (1960c) reported a rather complex control in cells of Chinese hamster involving both sex chromosomes and autosomes. Lima-de-Faria (1959) had suggested that the late replicating DNA was located in heterochromatin, but the two X-chromosomes of the Chinese hamster female were different in their replication. These observations suggested that the pattern of replication was a genetic trait, and like other genetic factors could mutate. If late replication was a characteristic of heterochromatin, this state was also mutable or subject to change at certain stages of differentiation (Lyons, 1961; Morishima *et al.*, 1962). Recent studies (Morishima *et al.*, 1962) have shown that the heterochromatic X-chromosome of the female human complement is late in completing its replication compared to its homologous mate and to most of the autosomes.

By the use of pulse-labeling techniques with thymidine-$H^3$ asynchronous labeling of sectors of chromosomes can be demonstrated in a variety of cells with relatively large chromosomes. However, of perhaps more significance is the observation that a large chromosome can be labeled at many sites simultaneously. For example, the Y-chromosome and the late replicating arm of an X-chromosome in the cells of Chinese hamster can be labeled at points along the whole length in 10 minutes of a total replication period of 3½ hours for these particular chromosomes (Taylor, 1960c). The late replicating X-chromosome of the human complement can likewise be labeled along the whole length in a 10-minute period although it, like the sex chromosomes of the hamster, can be labeled in isolated sectors when supplied with the isotope very near the end of the *S* period (Morishima *et al.*, 1962). This rapid labeling of DNA along the whole length of chromosomes was also noted in root cells of *Bellevalia* (Taylor, 1959a). These observations must mean that a large chromosome has several and probably a very large number of sites where DNA replication may be occurring simultaneously.

## VI. Chromosome Reproduction in Relation to DNA Replication and Protein Synthesis

### A. GENERAL COMMENTS

Chromosome reproduction involves replication of DNA and synthesis of proteins at least in higher organisms. Synthesis of RNA may be involved directly in the process but this has not been demonstrated so clearly as in the case of the other two polymers. DNA replication can be dissociated from protein and RNA synthesis in bacteria. By using a mutant of *Escherichia coli* that requires arginine, uracil, and thymine for growth, Maaløe and Hanawalt (1961) showed that DNA would increase by 40% or more without additional protein or RNA synthesis. This would be the amount formed if all nuclei in synthesis should finish but not begin another round of replication. By restoring arginine and uracil to allow a short period of protein and RNA synthesis, another replication could then be initiated. Cohen (1962) has found a change in state of the DNA after this transition period. If the DNA is isolated before protein synthesis is allowed to occur, in cells starved for arginine and uracil, the isolated DNA can be precipitated with streptomycin. After the period of protein synthesis and the initiation of a new cycle of replication, the DNA isolated from these cells is no longer precipitated by streptomycin. It is perhaps significant that the single-stranded DNA of $\phi$X174 is not precipitated by streptomycin. Cavalieri and Rosenberg (1961a) also reported a change in state of *E. coli* DNA when cells were starved for thymine for 45 minutes and then returned to a complete medium. DNA samples collected before and for a short period after restoring thymine had a molecular weight twice that of DNA isolated from cells about one-half hour after restoring thymine. They also found an associated change in the kinetics of degradation by DNase which they interpreted as a change from four-stranded DNA to two-stranded DNA. Further information on changes in state of DNA is given in Chapters III and IV.

### B. TYPES OF CHROMOSOMAL ABERRATIONS PRODUCED AT VARIOUS STAGES OF THE CELL CYCLE

Methods for detecting changes of chromosomes in relation to the duplication cycle are quite limited. One of the most revealing techniques is the study of changes produced by ionizing radiations. Soon after the mutagenic effects of X-rays had been demonstrated, there were attempts to determine the time of chromosome duplication by the type of aberrations observed after irradiation at different stages of the nuclear cycle in root cells (Mather and Stone, 1933). In early experiments most of the aberrations observed were the type expected if the chromosomes were broken before reproduction. Their conclusion was that chromosomes were single during most or all of interphase and reproduction occurred late in the division cycle. Their work was criticized by Huskins and Hunter (1935) on the basis that the stage irradiated could not be determined accurately, and that cells observed might not have been at the first division after treatment. Huskins and Hunter and also Mather (1937) improved on the design of the early experiments by selecting microspores of higher plants as experimental material. After meiosis the haploid microspores formed pass through a long interphase of several days to 2 or 3 weeks depending on the species and temperature of development. By irradiating in this interphase and analyzing cells at the first microspore division, they could be certain that the chromosomes were observed at the first division after treatment. However, Huskins and Hunter (1935) reported that the chromosomes of *Trillium* were already double when irradiated. Mather (1937) found that treatment at early interphase resulted in only chromosomal type aberrations, i.e., those produced before production of the two chromatids. At later stages of interphase single chromatids of a pair could be broken. Huskins and Hunter (1935) contended, along with others, that the degree of subdivision of a chromosome could not be revealed by irradiation because the passage of an ionizing particle could sever several subunits at once. Another possibility is that a chromosome broken before reproduction may remain open and react with other broken ends after reproduction.

While both of these possibilities may be realized in some instances, the reproduction of chromosomes determined by the types of aberrations produced shows a good correlation with other events that would be expected to occur during reproduction. Although both chromatids of a chromosome may be broken simultaneously (isolocus breaks) the breakage of one without the other being affected reveals a kind of reproduction or change of state. Although broken ends may not undergo reunion immediately, experimental evidence indicates that most breaks rejoin or the ends exchange with other ends produced by breaks close by within an hour or less in metabolically active cells (Wolff and Luippold, 1956a).

Sax (1941) studied the change in state of chromosomes of *Tradescantia* microspores by analyzing the types of aberrations in cells fixed at hourly intervals after X-radiation. The change from single chromosomes to double chromosomes occurred between 32 and 26 hours before metaphase. When the time of DNA synthesis was determined in these same cells by autoradiographic studies (Taylor, 1953) the two events were shown to coincide. Later a study of DNA synthesis by autoradiography and cytophotometric techniques at this stage and in the subsequent replication in the generative nucleus again showed a close correlation of the change to doubleness to radiation and the increase of DNA (Moses and Taylor, 1955). Thoday (1954) found a similar correlation in root cells of *Vicia*. The change in state did not appear to be abrupt but a gradual one spread over several hours. The experiments were not designed in such a way that they could reveal whether this difference was due to an asynchrony in the cell population or a gradual change of the chromosome complement of a single cell as its DNA was replicated. Recent experiments indicate that the change may be rather abrupt and may occur before DNA replication (Wolff, 1961; J. H. Taylor, unpublished). However, the abruptness of the change may vary with the cell type. In the cells studied DNA replication extends over a period of several hours and is sometimes asynchronous with respect to whole chromosomes or sectors of these. In *Vicia* and *Bellevalia* root cells, all chromosomes are labeled along most of their length nearly simultaneously, but in cells from other sources (*Crepis* roots, Chinese hamster fibroblasts, human leucocytes, and grasshopper gonial cells) individual chromosomes or sectors are asynchronous in their replication as indicated by labeling with thymidine-$H^3$ (Taylor, 1960c; Morishima *et al.*, 1962). A recent report by Chu and Giles (1961) shows that a chromatid exchange and a chromosome exchange can occur in a single cell of the human where the labeling mentioned above indicates asynchrony.

All of the information available indicates that the changed state, from that in which chromosomes act as single units in the formation of aberrations to the state in which they react as two units, is correlated in time with DNA replication. The change either occurs progressively during synthesis or rather abruptly preceding synthesis. The earlier discrepancies reported by Huskins and Hunter (1935) and by Mather (1937) are very likely the result of variations in the biological material. The time of synthesis of DNA has not been reported for the two genera used, *Trillium* and *Allium*. However, in *Tradescantia* the change and synthesis both occur in late interphase, while in *Tulbaghia* both DNA synthesis and the change in state occur in very early interphase (Taylor, 1958c).

In meiosis where chromosome reproduction has often been supposed to occur in the long prophase of the first meiotic division, the change in state from single to double occurs in interphase (Mitra, 1958; Sauerland, 1956). Likewise DNA replication and the synthesis of basic proteins of the nucleus are correlated with the change (Taylor, 1959c; De, 1961). The synthesis of basic proteins (histones) of chromosomes is apparently correlated with DNA replication in other cell types (Alfert, 1955; Bloch and Godman, 1955; Gall, 1959). In all of these studies the correlation of changes in response to irradiation and the synthesis of DNA and proteins lack the precision in timing for specific chromosomes or regions of chromosomes that would be desirable. Further discussion of a change in state of DNA in relation to its replication cycle will be found in Chapter IV and the last section of this chapter.

Another type of aberration induced by radiation at limited stages of the cell cycle are half-chromatid exchanges. They can be produced only in mitotic prophase (Sax, 1957) and in the stages after pachytene in meiosis (Mitra, 1958; Crouse, 1954). Although these connections are strong enough to lead to breakage of chromatids as they stretch in anaphase, the bridges do not persist through the subsequent interphase and usually do not become chromatid aberrations at the next division (Östergren and Wakonig, 1954). After anaphase the chromatids revert to a state in which they act as if composed of single axial elements in breakage and reunion rather than double as in the previous prophase. However, there is some evidence that they can become double to radiation breakage before DNA synthesis begins. See the last section of this chapter for a model that could behave in this fashion.

### C. HEALING AND REUNION OF BROKEN CHROMOSOMES

The synthesis of the major components of chromosomes is restricted to one-half or less of the cell cycle in many cells of higher organisms. Yet, the repair or reunion of broken chromosomes can occur at other stages of the cell cycle. Perhaps the clearest case is the breakage of chromosomes in the prophase stages in *Lilium* after a demonstration of DNA replication is no longer possible. These breaks produce chromatid bridges in abundance (Mitra, 1958). Other examples are the production of chromatid exchanges in pollen tubes and in the late interphase in roots. Breaks that are produced before DNA replication in the $G_1$ stage also undergo reunion or exchange before replication. This is indicated by the yield of chromosome type aberrations in some cells and perhaps more convincingly by the results of variations of intensity and fractionation of dose in radiation studies (Sax, 1939; Wolff and Luippold, 1956a). The yield of one-hit aberrations, i.e., single breaks, is linear with dose, but exchanges that require interaction between ends produced from separate breaks increase almost as the square of the dose. If the dose is given over a longer period so that most breaks are repaired before another is produced close enough for interaction, the curve of yield to dose approaches linearity. On the other hand, one-hit aberrations show no change with variations in rate, but are proportional to total dose. There is also a decrease of the two-hit type when the dose is given in two fractions separated by an interval long enough for nearly complete healing of lesions. Wolff and Luippold (1956b) obtained evidence for two kinds of aberrations in *Vicia* roots, a group that undergoes reunion in about 1 minute and a second group that rejoins within 1 to 2 hours. Since these experiments were carried out with *Vicia* seeds soaked for 18 to 24 hours in water, many of the cells are presumably in $G_1$. Exposing the cells to tritiated thymidine revealed no DNA synthesis (Wolff, 1960). These results indicate that broken chromosomes can rejoin before DNA replication. By exposing the roots to chloramphenicol (300 $\mu$g/ml) the incorporation of glycine-$C^{14}$ was reduced by 13–33%. When the chloramphenicol was given before or after a single dose of radiation it had no effect on aberrations, but when given for the period between two doses it prevented the usual split dose effect, i.e., the breaks from the first dose remained open and capable of exchanging with those produced by the second. Aureomycin had a similar effect, but penicillin did not prevent rejoining. On the basis of these experiments Wolff (1960) maintains that protein synthesis is necessary for chromosome reunion and supposes that the linear structure is maintained by protein. Since DNA replication could not be demonstrated during the period of rejoining, he assumes DNA is not involved.

However, recent experiments (Taylor *et al.*, 1962) indicate that DNA replication may be involved in chromosome reunion. Fluorodeoxyuridylate is a potent inhibitor of thymidylate synthetase. When fluorodeoxyuridine (FUDR) is presented to cells, many of them have enzymes for its phosphorylation. The nucleoside monophosphate binds with the enzyme that normally converts deoxyuridylate to thymidylate and blocks synthesis. In the absence of thymidine triphosphate, DNA replication ceases (see Chapter I). The root cells of *Vicia faba* are quite sensitive to the inhibitor. At a concentration of $10^{-7}$ *M* the rate of cell division drops to a low level within 6 hours, but the cells which arrive at division have intact chromosomes. At higher concentrations ($10^{-6}$ *M* and higher) the chromosomes show extensive damage. The damage appears about 3 hours after immersion of roots in the inhibitor. Lesions, which appear as non-staining gaps in the late prophase, metaphase, and anaphase chromosomes, are the first visual signs of damage. Some of the gaps are complete discontinuities of the chromatids and free fragments are the results. Frequently, but not regularly, both chromatids are affected at the same locus. Many of the last cells to arrive at division show an extensive shattering of all or most of the chromosomes. The lesions apparently occur in cells near the end of their DNA replication cycle when the enzyme, thymidylate synthetase, is quickly blocked. Such cells have no mechanism to prevent their progress through division even though their DNA replication is not complete. The lesions can be prevented by supplying exogenous thymidine along with the inhibitor. In addition, the lesions can be cured by giving the cells thymidine at least 1 hour before anaphase, when they are presumably in late interphase or early prophase. Bromodeoxyuridine can be substituted for thymidine in the healing of these lesions. Evidently, DNA replication, which requires all four nucleotides including thymidylate, is necessary for healing the lesions.

An interesting feature of the fragmentation is the absence of chromatid exchanges or any aberrations that will produce bridges at anaphase. This observation led to the idea that DNA replication might be required for reunion of chromosomes broken by radiation. A test of the hypothesis appeared possible because a concentration of $10^{-7}$ *M* stopped cell division and presumably DNA replication without yielding visible chromosome damage. By giving cells a dose of 25 r of X-irradiation 4 hours before fixing, anaphase bridges were produced in 6.4% of the cells. If the cells had been placed in $10^{-7}$ *M* FUDR for 30 minutes before irradiation and were grown in the same solution until fixed 4 hours later, only 1.1% of the anaphases had bridges. By supplying thymidine or bromodeoxyuridine ($10^{-4}$ *M*) to the irradiated cells treated with FUDR, the percentage of bridges could be restored to nearly the control level, 5.4 and 5.5%, respectively. From these experiments the hypothesis was proposed that reunion of broken chromosomes requires DNA replication. The process was visualized as involving two steps. The first would be a very rapid one that does not require high energy bonds, i.e., base pairing of two single polynucleotide chains (Fig. 6). The two chains of the DNA double helix were assumed to break at different points in many cases. The union by base pairing could be a very rapid process. The second step would presumably be the healing of the lesion by the growth of two antiparallel DNA chains, each copying the other. If the growth occurs by the action of the enzyme DNA polymerase (Chapter I) only chains with 3′ OH groups at the end can grow. If the breaks leave a 3′ phosphate group at the end, an enzyme might be required to chop the chain and prepare it for growth. In addition, about one half of the unions by base pairing would be expected to occur between non-growing ends, i.e., those ending in a 5′ phosphate or 5′ OH group. In these the shorter complementary chains ending in a 3′ OH might be expected to grow and bypass the original point of union of nongrowing ends. This could result in a reopening of the break and result in new exchanges or to a union

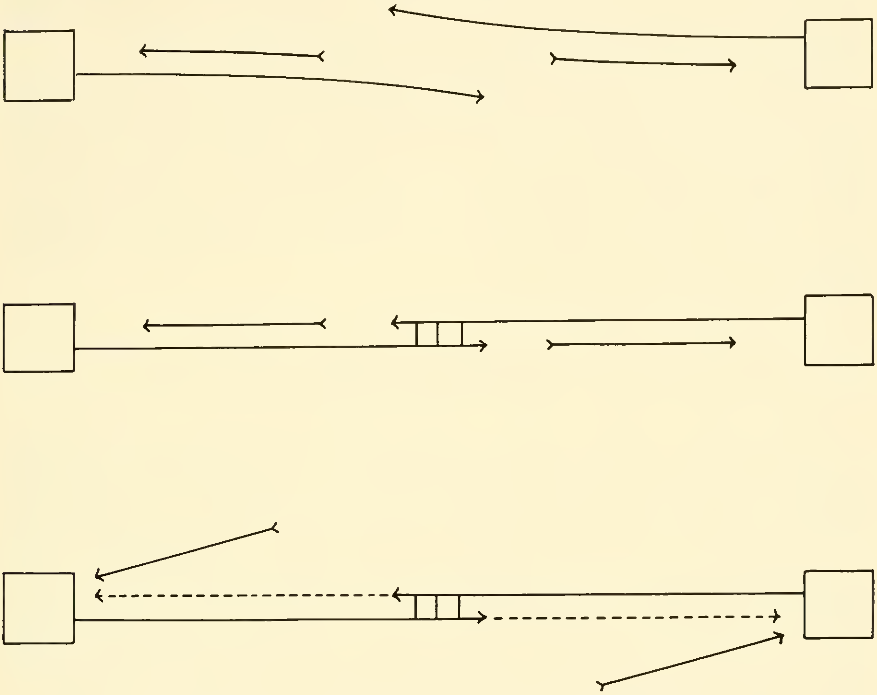


Fig. 6. Diagram of the steps involved in the reunion of broken chromosomes according to the hypothesis described in the text. Fragments of two replication units of different broken chromosomes are shown. The lines represent the subunits of chromatids, presumably single nucleotide chains. (From Taylor *et al.*, 1962.)

through the continued growth of the 3′ ends. One of the most interesting consequences of such a mechanism of reunion of broken ends would be the elimination of single strands of DNA over a part of the replicating unit. Perhaps more interesting is the idea that it would provide a mechanism for copy choice over short segments in combination with breakage and exchange between homologous chromatids. Such a mechanism of reunion could explain all of the types of aberrant recombination reported in ascomycetes (Olive, 1959; Kitani *et al.,* 1961; Case and Giles, 1958; Rizet *et al.*, 1961) and perhaps also in phages (Edgar, 1961).

The basis for the above hypothesis is somewhat weakened by subsequent studies of reunion of radiation-induced breakage. Chromatid exchanges apparently can occur in the presence of the block by FUDR. This would suggest that many of the breaks do not involve DNA or that another component, probably protein, is able to bring about reunion. Such a mechanism is suggested by the model of chromosome structure presented in the following section.

## VII. Molecular Models for Organization of DNA into Chromosomes

Models which allow us to explain data, to express more concretely certain concepts, and to plan and design future experiments are useful. However, models have the disadvantage that they tend to freeze our thinking in such a way that significant data may be neglected. When we are fully aware of such disadvantages models can be useful, even when the information for making them is limited. Several recent attempts have been made to construct models of a chromosome (Schwartz, 1955; Bloch, 1955; Taylor, 1957; Ris, 1957, 1961; Gall, 1958b; Freese, 1958; Taylor, 1958d; Steffensen, 1959; Stahl, 1961). All of these are useful in explaining some data and in crystallizing concepts, but all are deficient in other respects.

Any model of a chromosome based on the present information should be composed of one, or at most a few, DNA double helices that extend the length of a chromatid. These must function as two subunits during replication at all levels of organization. There is also considerable evidence that the chromatid functions as two subunits in non-reciprocal recombination (Olive, 1959) and in the induction and segregation of mutants induced by base analogs (Strelzoff, 1962). However, in chemically and radiation-induced breakage and exchange as well as reciprocal recombination and sister chromatid exchanges it acts as a single unit except during prophase. On the other hand, the pattern and frequency of sister chromatid exchanges indicate that, even though the exchanges are only between whole chromatids, each chromatid is composed of two unlike subunits (Taylor, 1958a). When all of the evidence is considered, the basis for this "twoness" would appear to be the complementary chains of the DNA double helix with some tertiary structure added (see below).

However, evidence is accumulating which indicates that chromosomal fibrils are double at many stages in the division cycle. For a review of the evidence from electron microscopy see papers by Ris (1957, 1961), Kaufmann *et al.* (1960), and Pappas and Brandt (1961). In interphase nuclei fibrils about 100 Å in diameter are seen in cells of many species. Often these are paired. The pairs of 100 Å fibrils could very well represent the two chromatids. However, the 100 Å fibrils which appear in spermatids before the chromosomes have been reproduced may still contain two DNA double helices, for according to Ris (1961) these change to 40 Å fibrils in most sperm nuclei. Although the decrease in diameter could result from a change in the protein with consequent variations in the pitch of a superhelix of the DNA double helix, it is possible that a significant characteristic of genetic material is being overlooked in a preoccupation with models based on single DNA double helices. The studies of Cavalieri and Rosenberg (1961a,b) on the physical properties of isolated DNA and Hall and Cavalieri's (1961) studies on the mass per unit length of DNA fibrils isolated from rapidly dividing *E. coli* also indicate that DNA from dividing cells is four-stranded while that from nondividing cells is two-stranded, i.e., is composed of single double helices. Is it possible to reconcile these observations with those which indicate that the two complementary chains of a single DNA double helix are the subunits for replication, mutation, and perhaps recombination?

A model based on a single DNA double helix which extends through the length of a large chromosome without interruption has some of the features required by our present data, but other necessary features would appear to be missing. One of these is the packing or folding of several centimeters of DNA for manipulation within the dimensions of a cell. The DNA or "chromosome" of phage T4 is more than 50 μ long when fully extended (Chapter III). The chromosome of *E. coli* would be perhaps 50 times this length (more than 2.5 mm). Some of the largest chromosomes might be more than a meter in length (Taylor, 1957). Even if these prove to be composed of several strands of DNA, they would still be several centimeters in length. The most useful models to visualize the folding and unwinding of such a long piece of DNA are those based on a suggestion by Freese (1958). The essential feature of the model is a DNA double helix with a regular sequence of linkers alternating in the two chains and located opposite a gap in the complementary chain. For modifications and other features of the model see papers by Taylor (1958d; 1959a,b; 1962) and Kellenberger (1960). Linkers have not yet been identified in analyses of DNA and indeed this can hardly be expected if they separate segments with molecular weights of several millions. Since such linkers can only be found by a search directed specifically to the problem and then probably only after improvement in present techniques, a little speculation on the nature of the linkers may be useful in the interim.

In addition to a capacity for folding and efficient packaging which might be provided for by irregularities such as odd bases, for example, methylaminopurine (Dunn and Smith, 1955 and 1958), a chromatid must have many growing points for replication and yet be closed to replication for long periods of function. The initiation and control of replication would appear to require operator sites to explain the observed control in order or sequence and to prevent more than one replication at each locus. To provide growing points breaks might be made by an enzyme, but to explain how the two subunits of a chromatid can remain intact or undergo only a few exchanges during replication the breaks must be produced in a very precise manner which would require regular structural sites for initiation of replication. To provide the necessary controls one may have to imagine several specific inducers for replication, the specificity of which need not reside in the type of linkers but in the neighboring nucleotides. In addition, a minimum of four types of linkers appears to be required. The first type of linker needed is one to join 5′ OH or 5′ phosphate groups at the ends of two polynucleotide chains with a reversal in polarity (Fig. 7). This will be referred to as a 5′ linker. Their complementary chains will be assumed to be joined by a 3′ linker. The chromatid is assumed to consist of a series of tandemly linked segments (replication units) of DNA with a 3′ OH group linked to a 3′ OH group and a 5′ end linked to another 5′ end at each operator site. Any irregularity in single chains of the DNA double helix would appear to be inadequate, for these would tend to be redistributed by inversions or other chromosomal aberrations in such a way as to prevent the subunits from maintaining their alternating sequence of linkers. Without such a sequence there would be numerous mistakes in sorting during replication and the subunits would not maintain their integrity. With linkers of the type proposed above, chromosomal aberrations might change the length of segments but would not destroy the alternating sequence of 3′ and 5′ linkers.

The 5′ linkers would always be closed and in non-replicating DNA the 3′ linkers would also be closed. One of the first steps preparatory to

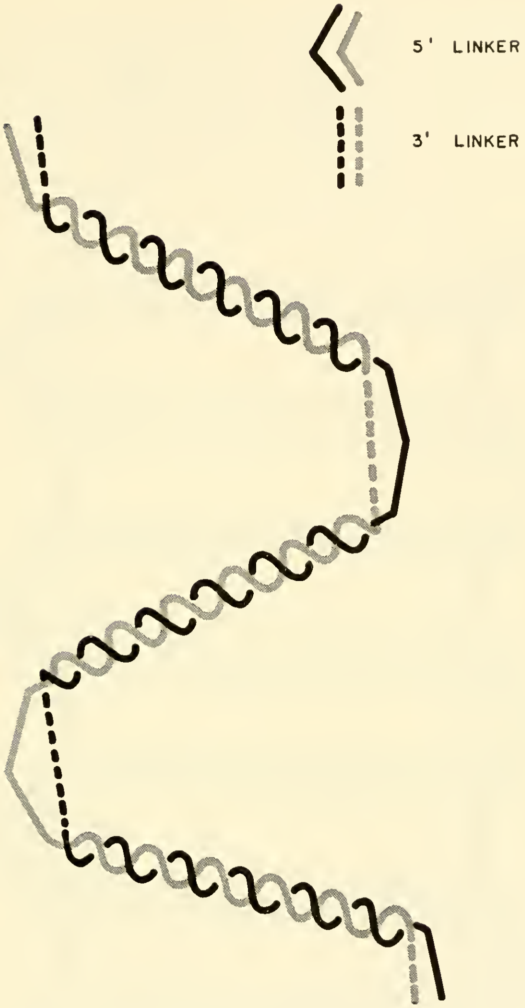


FIG. 7. A portion of a chromosome preceding replication. The plectonemically coiled strands represent molecules of DNA joined in tandem array by means of special sites or linkers in each polynucleotide chain. These linkers are assumed to occur in pairs at sites in which there is a reversal in polarity so that two 3′ and two 5′ ends of chains are linked by 3′ linkers and 5′ linkers, respectively.

replication would be the enzymatic opening of 3′ linkers. However, a necessary preliminary step would be the establishment of H linkers (Fig. 8) which are assumed to stabilize the chromatid by holding it in a folded ladder-like arrangement (Taylor, 1962). In the model the H linkers are shown as some type of polymer joining alternate 5′ linkers along two axes. These would represent the half-chromatids, but during

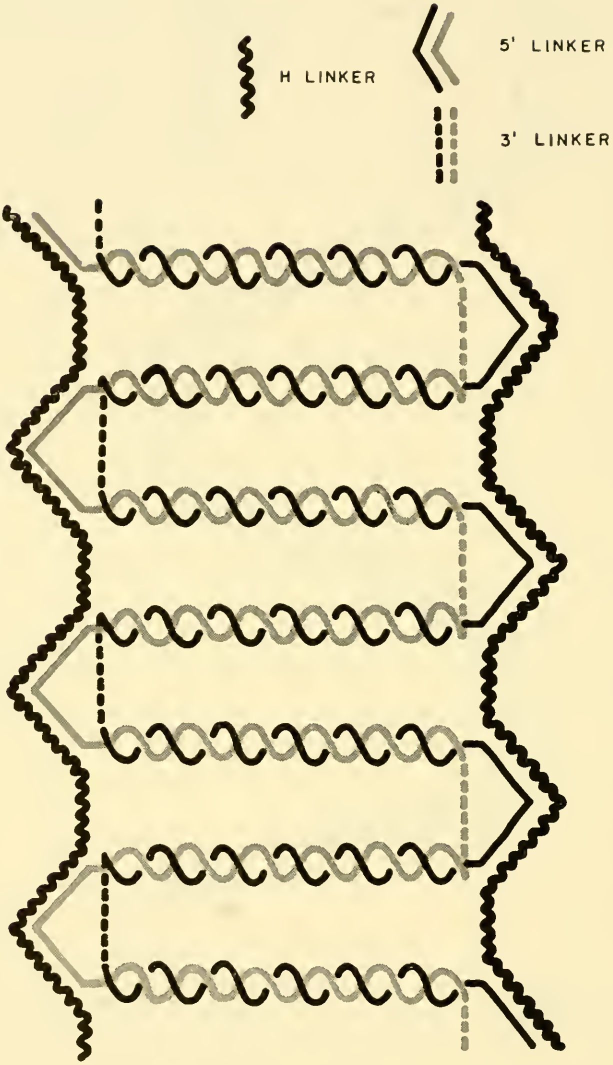


Fig. 8. Before or during replication the chromosome is assumed to be stabilized by the addition of H linkers to form a double axis, each of which with its attached DNA chains will become a chromatid by the time replication is complete. The stabilizing material is represented as a polymer attached to alternate 5′ linkers. Although shown as a continuous chain for simplicity in illustration, it probably would consist of many small subunits attached in some way as yet unknown.

and following replication would form part of the linear axis of the two sister chromatids. A new set would be formed on each chromatid at prophase when it became double, i.e., when half-chromatid bridges can be induced. In rapidly dividing cells these linkers might be closed most of the time, but in cells where the chromosome acts as a single unit in breakage and reunion these linkers would have to be assumed to be open or rather labile. Since protein synthesis appears to play a role in the reunion of chromosomes, the H linkers for the present will be assumed to consist of polypeptides.

As soon as the H linkers are stabilized the chromosome should be capable of acting as a double structure in breakage and exchange. In *Vicia* this appears to occur for all chromosomes before DNA replication begins or at the beginning of the *S* period (see Section VI). In other species where there is a marked asynchrony of replication, the stabilization would be expected to occur in some chromosomes or sectors of chromosomes before its completion in others.

Replication is assumed to begin at a pair of the 3′ ends, perhaps by the insertion of a 5′ linker and the closure of the 3′ ends by a new and different 3′ linker not susceptible to the enzyme which is operating to open the regular 3′ linkers. Since this 3′ linker is necessary only during the replication cycle, we will refer to it as the 3′R linker. It must be changed to a regular 3′ linker before another replication can occur. Some evidence suggests that the change may occur during anaphase or telophase (see below). To allow rotation and untwisting of the DNA helices the 3′ linkers should be open on the side opposite the insertion of the new 5′ linkers. If there were an out-of-phase replication of the two subunits of a chromatid, unwinding would appear to be facilitated. However, in an organism where the rate of DNA replication may be a factor in limiting growth, *E. coli* for example, the cell must have evolved some control so that both subunits can be replicating simultaneously.

When two adjacent chains finish replication we will assume that the 3′ ends of the two new chains are now closed with a 3′R linker (Fig. 9). The result would be a replicated package consisting of a pair of double helices attached to one axis of the original chromatid. The complementary chains attached to the other axis would now replicate and form other pairs of helices attached to a single axis. When both chromatids were complete, each would consist of a series of the four-stranded packets of DNA attached to a linear axis with a length that might very well be manipulated in the dimensions available in a cell (Fig. 10). For example, in *E. coli* if the units of replication had a molecular weight of $2.4 \times 10^6$ (Cavalieri and Rosenberg, 1961a), and were attached 40 Å apart along the axis, the over-all length of the chromosome would be only

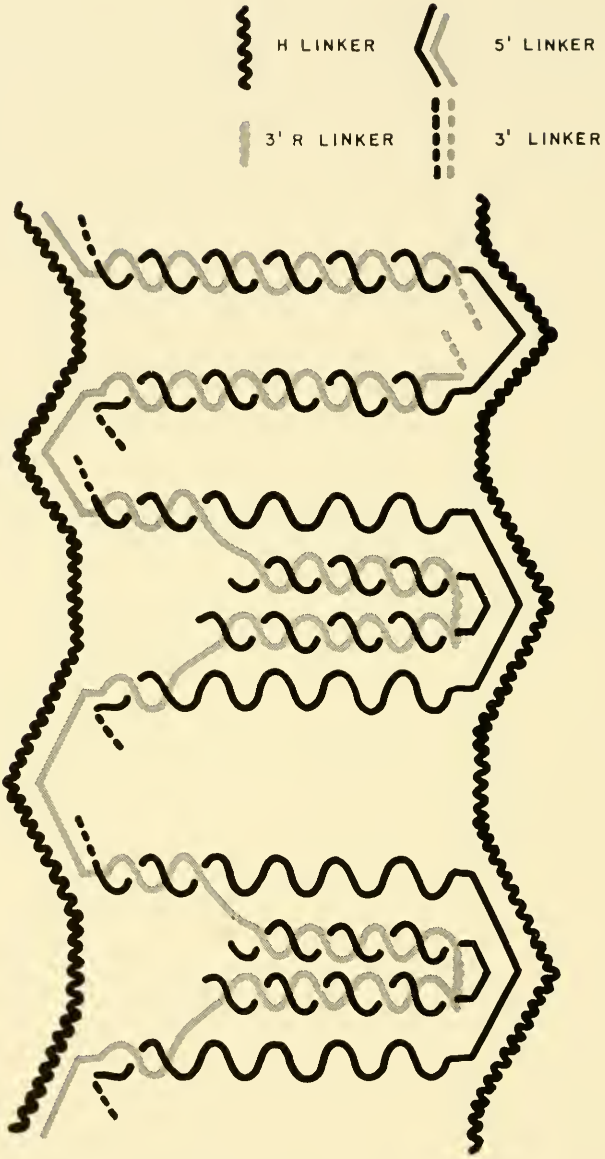


Fig. 9. During replication the 3′ linkers are assumed to be opened by appropriate enzymes. The two subunits of the chromosome consisting of the alternate 5′ linkers and the H linkers with the attached DNA chains are then held together only by the hydrogen bonds in the DNA. Two chains of adjacent molecules are assumed to begin replication by the insertion of a new 5′ linker.

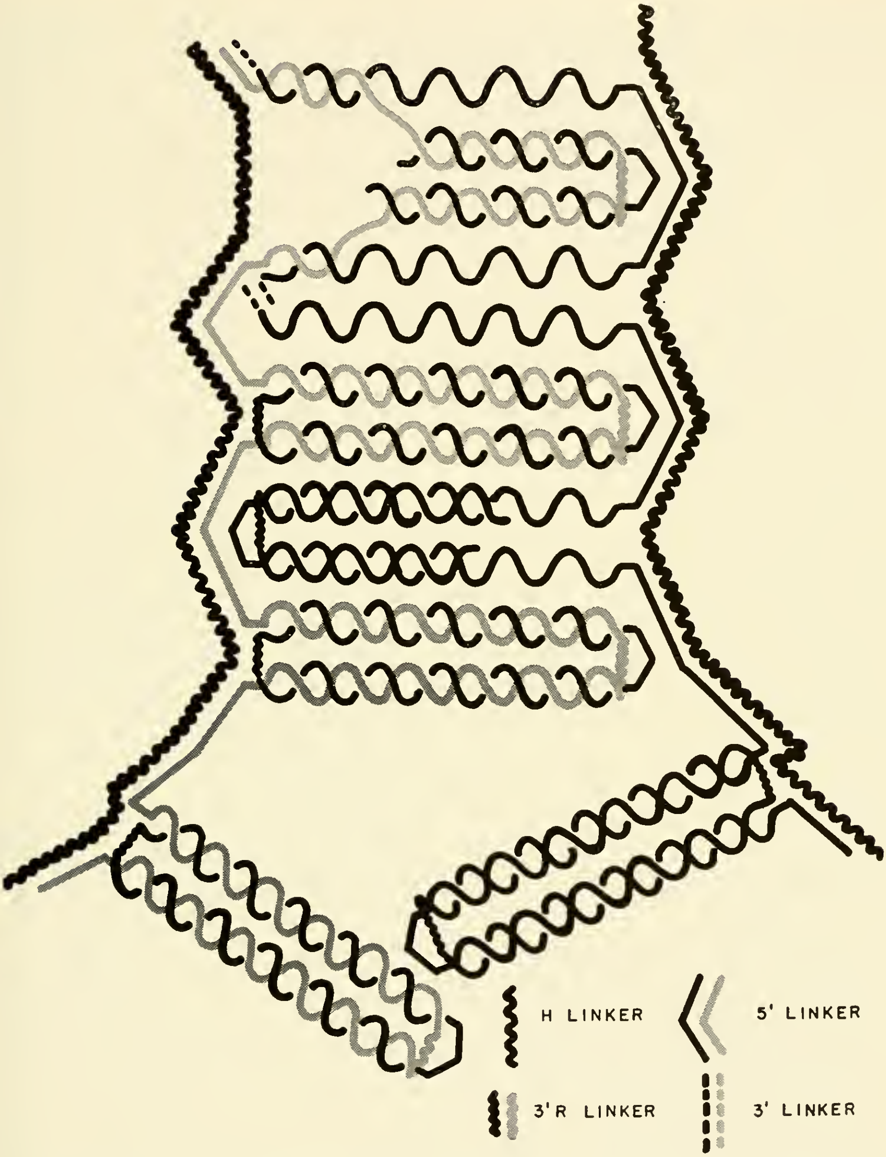


FIG. 10. As each pair of chains begins replication, a new 3′R linker is assumed to be inserted which is resistant to the attack of the enzyme that opens the regular 3′ linkers. When the pairs of chains finish replication their growing 3′ ends are assumed to be closed by the addition of another 3′R linker. The two new chromatids would consist of these pairs of DNA molecules attached to a single axial element. Another axial element (half-chromatid) would be added to each during the succeeding late interphase or prophase.

3 or 4 $\mu$. A large chromosome would likewise be of a size so that one or two sets of coils in the ribbon formed by the ladder-like arrangement of DNA helices would produce the structure seen with the light microscope (Taylor, 1959d; Taylor *et al.*, 1957).

In $G_2$ each chromatid should consist of a single axial element made of alternating H linkers and 5′ linkers. DNA extracted at this stage should be a very poor primer for DNA polymerase unless the chains in replicated units were broken to provide growing points. By some stage of prophase when half-chromatid exchanges can be induced the new set of H linkers may be assumed to have formed and become stabilized so that each chromatid has two axial elements. By late anaphase or early telophase some chromatids become separable into two subunits (in most chromosomes which are twisted the plectonemic coiling of the two axial elements, half-chromatids, prevents complete separation). The separation can be brought about by agents which would be expected to break hydrogen bonds and perhaps denature proteins, but would not be expected to break covalent bonds. This is the basis for the idea that the 3′R linkers may be opened at this time and presumably replaced by regular 3′ linkers at a later stage. If that is the correct interpretation, the DNA might now be converted to a potentially active primer if the enzyme for opening the regular 3′ linkers were available. However, there is no assurance that this is the time of the change even if the model is correct. We know that chromosomes may undergo duplication without passing through anaphase. However, in the cases studied there is an endomitotic reorganization of some sort before a second duplication occurs.

The model can be used to account for most or all of the presently known properties of chromosomes. It is based on morphological and biochemical evidence drawn from a variety of sources where information in all cases is at best fragmentary. The morphological counterparts are the lampbrush chromosomes and the structures found by Pappas and Brandt (1960; and personal communication) in the interphase nucleus of amoeba (Fig. 11). The structures consist of remarkably well-preserved helical filaments clustered as if attached by one end to an axial element. I interpret these to be $G_2$ chromosomes of the type shown in the model (Fig. 10). The helical filaments have a gyre diameter of about 300–350 Å and an over-all length of 3000–4000 Å. The fiber wound into this helix is about 120 Å in diameter and has a periodicity indicating that a smaller fiber of about 40 Å is wound into a helix to form this 120-Å fiber. At certain loci the coiled filaments are separated enough to show that some at least are double, i.e., consists of two strands of 15–20 Å (Fig. 12). Measurements indicate that the ultimate helical filaments would have a contour length of 3–4 $\mu$. Each of these two associated

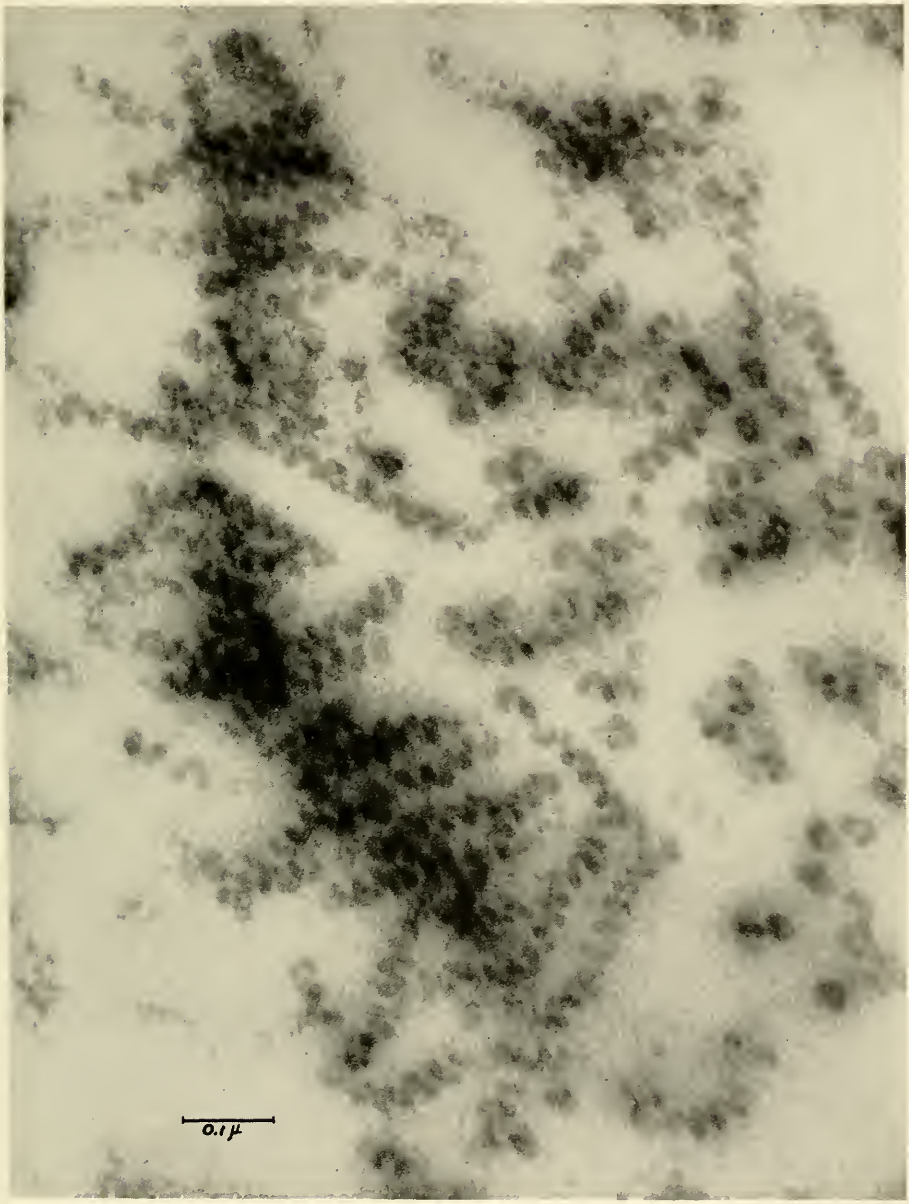


FIG. 11. Helical structures revealed by the electron microscope in interphase nuclei of *Amoeba proteus*. These helices appear to be attached along an axis which is assumed to represent a portion of a chromosome after replication according to the hypothesis in the text. (Photograph from Pappas and Brandt, 1960.)

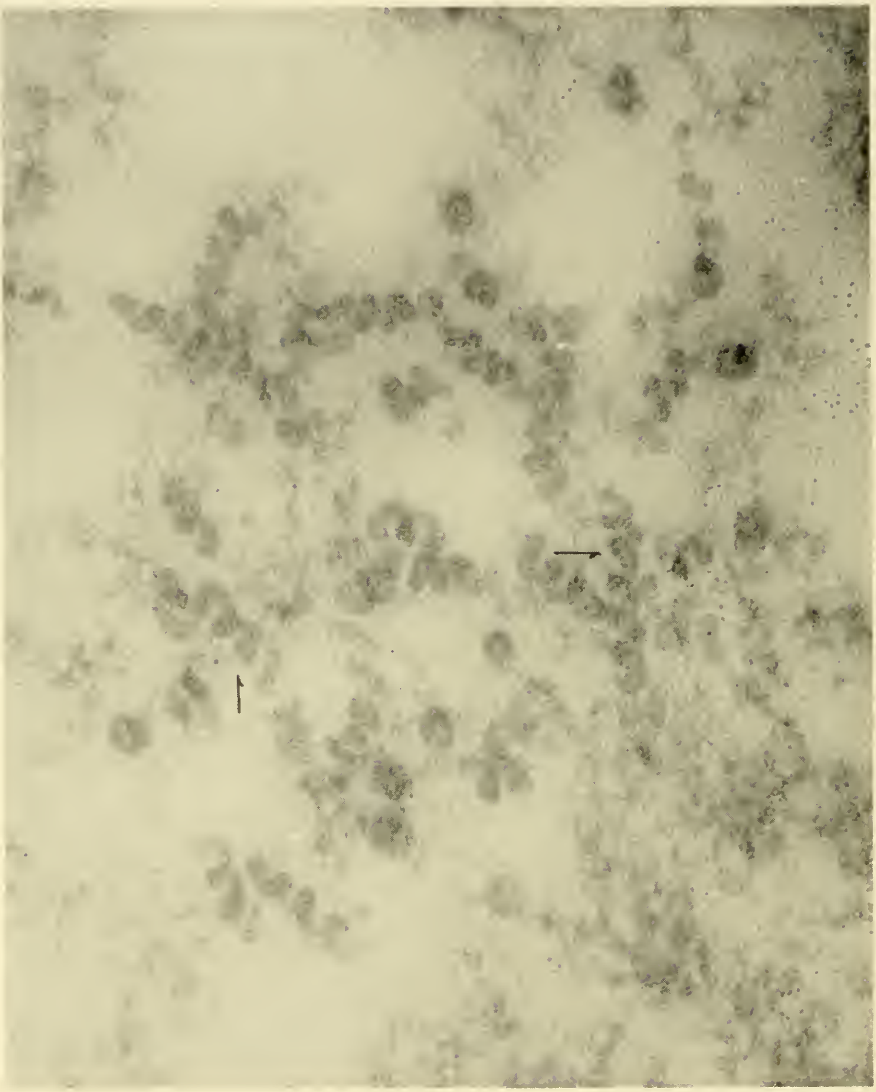


FIG. 12. Higher magnification of the helical structures shown in Fig. 11. At places the helices (*arrows*) can be seen to be double and the fibrils wound into the large helix have a periodicity of density indicating that this fibril is formed from a helically coiled ultimate fibril of about 30 Å. (Original photograph contributed by George D. Pappas and Phillip W. Brandt.)

strands (DNA double helices perhaps with some associated histone) would have a molecular weight of 6–8 × $10^6$ (Taylor, 1960b) and presumably would represent the replicating unit in amoeba.

The lampbrush chromosomes seen in meiosis may be a similar type of structure except that the axial element has a much greater density to the electron beam (Moses, 1956, 1960). Recently Bullivant (1962) has demonstrated helically coiled filaments in metaphase chromosomes of mammalian cells. The gyre diameter is about 75 Å and these helices occasionally appear to be coiled into a super helix comparable to those found in the amoeba. The chromosomes in amphibian oöcytes have loops, the axis of which appear to be single DNA double helices (Miller, 1962). These could be the pairs of replication units completely uncoiled which have opened up into rings. When stretched along the axis they break (Callan, 1956) at what would be the 3′R linkers in the model.

The pairs of replication units in the model could also correspond to the four-stranded DNA units reported by Cavalieri and Rosenberg (1961a,b). If a chromatid built according to the model is stripped of its protein (including the H linkers) during $G_2$, its weak points will be the 5′ linkers. The separated units will consist of rings in which one polynucleotide chain is interrupted with two free 5′ ends. If one assumes some type of biunial bonds the two double helices may remain paired. In rapidly dividing cells with a very short $G_1$ stage most of the DNA might exist in this state. Upon heating the rings could open and allow unwinding of the polynucleotide chains, with a decrease of the molecular weight by a factor of two and the separation of chains of different densities. Thus the observations of Meselson and Stahl (1958) on the separation of hybrid molecules into single chains could apply to four-stranded DNA as well as two-stranded DNA.

The model has at certain stages some characteristics of a model composed of tandemly linked rings, a model which has been suggested by Stahl (1961) to explain recombination events in higher organisms. However, at other stages the model acts like a side-chain model for purposes of genetic recombination (Taylor, 1962).

Very little can be said at this time about the nature of the hypothetical 3′, 3′R, and 5′ linkers. The H linkers have been assumed tentatively to be peptides, but it now appears that the other linkers cannot be peptides. At the chromatid level the best evidence comes from the report (Callan and McGregor, 1958; Gall, 1958b) that proteases and RNase do not sever the loops of the lampbrush chromosomes while DNase quickly disrupts them. In addition, isolated deproteinized DNA is not significantly degraded by proteases. A recent report by Bendich and Rosenkranz (1962) indicates that sperm DNA contains about one phosphoserine residues per thousand nucleotides. This suggests that the regular 3′ linkers or perhaps the 3′R linkers might consist of two phosphoserine residues coupled to the terminal nucleoside of DNA chains by an ester linkage in the same way that amino acids are coupled to transfer RNA. A diphosphate bridge could then couple two chains with a reversal in polarity. The linkage might represent an activated state ready for the initiation of replication or for a change of the linker if it should prove to be the 3′R linker. The 5′ linker should be formed by covalent bonds which might be as stable as the phosphodiester linkages along the polynucleotide chains. Although there are several possible types of molecules that could serve as linkers, it would be premature to make further guesses at this time.

Although the model presented is designed primarily to explain the replication and sorting of subunits in chromosomes, the traffic control problem in template function might be mentioned. How is a chromosome able to open for replication without functioning as a template for RNA synthesis as appears to be true for the heteropycnotic X-chromosome and perhaps other such genetically inactivated regions on autosomes? Or what is more common, how does it open for RNA replication without serving as a template for DNA synthesis when all the enzymes and precursors for replication are operating at other sites in the nucleus? The cell could solve this problem by having special operator sites which are opened for RNA replication, but which could not serve as starting ends for DNA polymerase. Let us suppose that a DNA polymerase is available to repair random breaks in DNA by adding to free 3′ OH ends and that the same or a similar polymerase can begin to copy at the proposed operator sites for DNA replication. However, the special sites opened for RNA synthesis would not be available as starting loci for these DNA polymerases because of the occurrence of some odd base, or other end group, for example. Such a mechanism would seem to be necessary to simplify the problem of traffic control in the operation of a template which can serve either for RNA synthesis or DNA synthesis. The principal difference would be that DNA replication requires opening and unwinding of the two chains of the DNA. RNA replication either can occur without strand separation or the chains can open at special sites and then close again without destroying the integrity of the DNA.

## REFERENCES

Alfert, M. (1955). *Symposium on Fine Structure of Cells, Leiden, 1954. Intern. Union Biol. Sci. Publ., Ser. B*, **21**, 157. P. Noordhoff, Groningen, Netherlands.

Avery, O. T., MacLeod, C. M., and McCarty, M. (1944). *J. Exptl. Med.* **79**, 137.

Bendich, A., and Rosenkranz, H. S. (1962). *In* "Progress in Nucleic Acid Research" (W. E. Cohn and J. N. Davidson, eds.), Vol. I. Academic Press, New York. In press.

Bloch, D. (1955). *Proc. Natl. Acad. Sci. Wash.* **41**, 1058.

Bloch, D., and Godman, C. G. (1955). *J. Biophys. Biochem. Cytol.* **1**, 17.

Boivin, A., Vendrely, R., and Vendrely, C. (1948). *Compt. rend. acad. sci.* **226**, 1061.

Bullivant, S. (1962). Personal communication; State of New York, Dept. of Health, Albany, New York.

Callan, H. G. (1955). *Symposium on Fine Structure of Cells, Leiden, 1954. Intern. Union Biol. Sci. Publ., Ser. B,* **21**, 89. P. Noordhoff, Groningen, Netherlands.

Callan, H. G., and McGregor, H. C. (1958). *Nature* **181**, 1479.

Case, M. E., and Giles, N. H. (1958). *Cold Spring Harbor Symposia Quant. Biol.* **23**, 119.

Cavalieri, L. F., and Rosenberg, B. H. (1961a). *Biophysics J.* **1**, 337.

Cavalieri, L. F., and Rosenberg, B. H. (1961b). *Biophys. J.* **1**, 317.

Cavalieri, L. F., and Rosenberg, B. H. (1961c). *Biophys. J.* **1**, 323.

Chu, E., and Giles, N. H. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 830.

Chun, E. H. L., and Littlefield, J. W. (1961). *J. Mol. Biol.* **3**, 668.

Cleveland, L. R. (1949). *Trans. Am. Phil. Soc.* **39**, 1.

Cohen, S. S. (1962). Personal communication; Dept. of Bioch. Univ. Penn. School of Med., Philadelphia.

Creighton, H., and McClintock, B. (1931). *Proc. Natl. Acad. Sci. U.S.* **17**, 492.

Crouse, H. V. (1954). *Science* **119**, 485.

Davison, P. F., and Levinthal, C. (1961). *Ann. Rev. Biochem.* **30**, 641.

De, D. N. (1961). *The Nucleus* **4**, 1.

Delbrück, M., and Stent, G. S. (1957). *In* "The Chemical Basis of Heredity" (W. D. McElroy and B. Glass, eds.), pp. 699–736. Johns Hopkins Press, Baltimore.

Djordjenic, B., and Szybalski, W. (1960). *J. Exptl. Med.* **112**, 509.

Doty, P. (1960). *Harvey Lecture* **55**, 103–139.

Dunn, D. B., and Smith, J. D. (1955). *Biochem. J.* **60**, xvii (abstract).

Dunn, D. B., and Smith, J. D. (1958). *Biochem. J.* **68**, 627.

Edgar, R. S. (1961). *Virology* **13**, 1.

Forro, F., and Wertheimer, S. A. (1960). *Biochem. et Biophys. Acta* **40**, 9.

Freese, E. (1958). *Cold Spring Harbor Symposia Quant. Biol.* **23**, 13.

Gall, J. G. (1956). *In* "Mutation," Brookhaven Symposia in Biology No. 8, pp. 17–32. Brookhaven National Laboratory, Upton, New York.

Gall, J. G. (1958a). Personal communication; Dept. of Zool., Univ. Minn., Minneapolis.

Gall, J. G. (1958b). *In* "The Chemical Basis of Development" (W. D. McElroy and B. Glass, eds.), pp. 103–135. Johns Hopkins Press, Baltimore.

Gall, J. G. (1959). *J. Biophys. Biochem. Cytol.* **5**, 295.

Gall, J. G. (1962). Personal communication; Dept. Zool., Univ. Minn., Minneapolis.

Geiduschek, E. P. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 950.

Hall, C. E., and Cavalieri, L. F. (1961). *J. Biophys. Biochem. Cytol.* **10**, 347.

Hershey, A. D., and Chase, M. C. (1952). *J. Gen. Physiol.* **36**, 39.

Hughes-Schrader, S. (1940). *Biol. Bull.* **78**, 312.

Huskins, C. L., and Hunter, A. W. S. (1935). *Proc. Roy. Soc.* **B117**, 22.

Kaufmann, B. P., Gay, H., and McDonald, M. R. (1960). *Intern. Rev. Cytol.* **9**, 77.

Kellenberger, E. (1960). *10th Symposium Soc. Gen. Microbiol.,* pp. 39–66.

Kellenberger, G., Zichichi, M. L., and Weigle, J. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 869.

Kitani, Y., Olive, L. S., and El-Ani, A. S. (1961). *Science* **134**, 668.

Kozinski, A. W. (1961). *Virology* **13**, 124.

LaCour, L. F., and Pele, S. R. (1958). *Nature* **182**, 506.

Levinthal, C. (1956). *Proc. Natl. Acad. Sci. U.S.* **42**, 394

Levinthal, C., and Crane, H. R. (1956). *Proc. Natl. Acad. Sci. U.S.* **42**, 436.

Lima-de-Faria, A. (1959). *J. Biophys. Biochem. Cytol.* **6**, 457.

Lima-de-Faria, A., Sarvella, P., and Morris, R. (1959). *Hereditas* **45**, 467.

Lyons, M. E. (1961). *Nature* **190**, 372.

Maaloe, O., and Hanawalt, P. C. (1961). *J. Mol. Biol.* **3**, 144.

McClintock, B. (1938). *Genetics* **23**, 315.

Mather, K. (1937). *Proc. Roy. Soc.* **B124**, 97.

Mather, K., and Stone, L. H. A. (1933). *J. Genet.* **28**, 1.

Marmur, J., and Doty, P. (1961). *J. Mol. Biol.* **3**, 585.

Meselson, M. (1960). *In* "The Cell Nucleus" (J. S. Mitchell, ed.), pp. 240–245. Academic Press, New York.

Meselson, M., and Stahl, F. W. (1958). *Proc. Natl. Sci. U.S.* **44**, 671.

Meselson, M., and Weigle, J. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 857.

Meselson, M., Stahl, F. W., and Vinograd, J. (1957). *Proc. Natl. Acad. Sci. U.S.* **43**, 581.

Miller, I. R. (1961). *J. Mol. Biol.* **3**, 229.

Miller, O. L. (1962). Personal communication; Biol. Div., Oak Ridge Natl. Lab., Oak Ridge, Tenn.

Mirsky, A. E., and Ris, H. (1949). *Nature* **163**, 666.

Mitra, S. (1958). *Genetics* **43**, 771.

Morishima, A., Grumbach, M. M., and Taylor, J. H. (1962). *Proc. Natl. Acad. Sci. U.S.* **48**, 756.

Moses, M. J. (1956). *J. Biophys. Biochem. Cytol.* **2**, 215.

Moses, M. J. (1960). *4th Intern. Kongr. Elektronenmikroskopie, Berlin, 1958,* pp. 199–211.

Moses, M. J., and Taylor, J. H. (1955). *Exptl. Cell Res.* **9**, 474.

Östergren, G., and Wakonig, T. (1954). *Botan. Notiser* **4**, 357.

Olive, L. S. (1959). *Proc. Natl. Acad. Sci. U.S.* **45**, 727.

Pappas, G. D., and Brandt, P. W. (1960). *4th Intern. Kongr. Elektronenmikroskopie, Berlin, 1958,* pp. 244–246.

Plaut, W. S., and Mazia, D. (1956). *J. Biophys. Biochem. Cytol.* **2**, 573.

Ris, H. (1957). *In* "The Chemical Basis of Heredity" (W. D. McElroy and B. Glass, eds.), pp. 23–62. Johns Hopkins Press, Baltimore.

Ris, H. (1961). *Can. J. Genet. Cytol.* **3**, 95.

Rizet, G., Lissouba, P., and Moussear, J. (1961). *Bull. soc. franç. physiol. végétale* **6**, 175.

Sauerland, H. (1956). *Chromosoma* **7**, 627.

Sax, K. (1941). *Cold Spring Harbor Symposia Quant. Biol.* **9**, 93.

Sax, K. (1939). *Proc. Natl. Acad. Sci. U.S.* **25**, 225.

Sax, K. (1957). *Quart. Rev. Biol.* **32**, 15.

Schildkraut, C. L., Marmur, J., and Doty, P. (1961). *J. Mol. Biol.* **3**, 595.

Schwartz, D. (1955). *J. Cellular Comp. Physiol.* **45**, Suppl. 2, 171.

Simon, E. H. (1961). *J. Mol. Biol.* **3**, 101.

Stahl, F. W. (1961). *J. chim. phys.* **58**, 1072–1077.

Steffensen, D. (1959). "Structure and Function of Genetic Elements," Brookhaven Symposia in Biology No. 12, pp. 103–124. Brookhaven National Laboratory, Upton, New York.

Stern, C. (1931). *Biol. Zentr.* **51**, 547.
Strelzoff, E. (1962). *Zeitsch. Vererb.* **93**, 301.
Sueoka, N. (1960). *Proc. Natl. Acad. Sci. U.S.* **46**, 83.
Taylor, J. H. (1953). *Exptl. Cell Research* **4**, 169.
Taylor, J. H. (1957). *Am. Naturalist* **91**, 209.
Taylor, J. H. (1958a). *Genetics* **43**, 515.
Taylor, J. H. (1958b). *Exptl. Cell Research* **15**, 350.
Taylor, J. H. (1958c). *Am. J. Botany* **45**, 123.
Taylor, J. H. (1958d). *Sci. American* **198**, 36.
Taylor, J. H. (1959a). *Proc. 10th Intern. Congr. Genet., Montreal, 1958,* pp. 63–78.
Taylor, J. H. (1959b). *Oklahoma Conf. Radioisotopes in Agr., 1959, U.S. Atomic Energy Comm. TID-7578,* pp. 123–129.
Taylor, J. H. (1959c). *Am. J. Botany* **46**, 477.
Taylor, J. H. (1959d). *In* "Symposium on Molecular Biology" (R. E. Zirkle, ed.), pp. 304–320. Univ. Chicago Press, Chicago.
Taylor, J. H. (1960a). *Advances in Biol. and Med. Phys.* **7**, 107.
Taylor, J. H. (1960b). *In* "Cell Physiology of Neoplasia" (M. D. Anderson Hosp. and Tumor Inst.), pp. 547–575. Univ. Texas Press, Austin.
Taylor, J. H. (1960c). *J. Biophys. Biochem. Cytol.* **7**, 455.
Taylor, J. H. (1962). *Intern. Rev. Cytol.* **13**, 39.
Taylor, J. H., Woods, P. S., and Hughes, W. T. (1957). *Proc. Natl. Acad. Sci. U.S.* **43**, 122.
Taylor, J. H., Haut, W. F., and Tung, J. (1962). *Proc. Natl. Acad. Sci. U.S.* **48**, 190.
Thoday, J. M. (1954). *New Phytologist* **53**, 511.
Watson, J. D., and Crick, F. H. C. (1953a). *Nature* **171**, 737.
Watson, J. D., and Crick, F. H. C. (1953b). *Nature* **171**, 964.
Wolff, S. (1960). *Am. Naturalist* **44**, 85.
Wolff, S. (1961). *Radiation Research* **14**, 517 (abstract).
Wolff, S., and Luippold, H. E. (1956a). *In* "Progress in Radiobiology" (J. S. Mitchell, B. E. Holmes, and C. L. Smith, eds.), pp. 217–221. Oliver and Boyd, Edinburgh
Wolff, S., and Luippold, H. E. (1956b). *Proc. Natl. Acad. Sci. U.S.* **42**, 510.
Woods, P. S., and Schairer, M. V. (1959). *Nature* **183**, 303.

## *Chapter* III

# The Organization of DNA in Bacteriophage and Bacteria

CHARLES A. THOMAS, JR.

*The Johns Hopkins University,*
*Baltimore, Maryland*

I. Bacteriophage DNA . . . . . . . . . . . . . . . . . 113
  A. The Organization of DNA in the Intact Phage Particles . . . . . 114
  B. The Structure of the Liberated Phage DNA Molecule . . . . . 121
  C. The Genetic Structure of Phage . . . . . . . . . . . . . 131
  D. The Distribution of Parental Phage DNA to Progeny Phage . . . . 138
II. Bacterial DNA . . . . . . . . . . . . . . . . . . 141
  A. Cytology of Bacteria . . . . . . . . . . . . . . . 142
  B. Growth and Transfer Experiments . . . . . . . . . . . 142
  C. The Structure of DNA Liberated from Bacteria . . . . . . . 145
  D. The Genetic Structure of Bacteria . . . . . . . . . . . 145
  E. Conclusions . . . . . . . . . . . . . . . . . 148
  Acknowledgments . . . . . . . . . . . . . . . . . 148
  References . . . . . . . . . . . . . . . . . . 148

## I. BACTERIOPHAGE DNA

The extensive development of the field of bacteriophage genetics has led to an increasing interest in the structure and organization of bacteriophage DNA. Since bacteriophage are the smallest biological units which are susceptible to thorough genetic analysis, and since their DNA is best described in molecular or chemical terms, a better understanding of the structure of phage DNA promises to lead to a chemical interpretation of genetic recombination, replication, and transcription. In addition to the genetic functions of DNA, this molecule has the important ability to assume a folded, compact configuration such as found in chromosomes, spores, and sperm heads. A form of this tertiary structure of DNA is found in the compact arrangement of the DNA inside bacteriophage. One may hope that the study of the organization of DNA in the phage head will lead to an understanding of the principles and requirements of this folding process.

It is convenient to divide these studies into those performed on the intact virus particles, and those on the liberated DNA molecules.

## A. THE ORGANIZATION OF DNA IN THE INTACT PHAGE PARTICLES

### 1. *Studies on Partially Oriented Preparations*

If the DNA in the phage head were organized in a special manner, it would be likely that this would be revealed by a preferential alignment of regions of the molecule with respect to the long axis of the phage.

When a concentrated suspension of anisometric particles is subjected to a stress, the long axes of the particles are preferentially aligned in the direction of the stress. It was by this technique that tobacco mosaic virus (TMW) was first aligned (Bernal and Fankuchen, 1941). By allowing a drop of concentrated T2 phage suspension to dry under a cover slip on a microscope slide, Bendet, Goldstein, and Lauffer (1960) found that there was a *negative* birefringence with respect to the direction of flow, while TMV displayed a positive birefringence in agreement with the studies of others. Since the birefringence of TMV is known to be positive (Franklin, 1955), while the birefringence of DNA is negative (Wilkins, Gosling, and Seeds, 1951), it was concluded that there is a preferential alignment of DNA molecules in the direction of the long axis of the phage particle. As the authors point out, there is a great danger that DNA is ejected as a result of shearing the concentrated suspension and certain precautions were taken to make this less likely.

A more quantitative study of this same phenomenon was done by Gellert (1961, 1962) employing flow-birefringence techniques. These experiments revealed that the rotary diffusion constant, and hence the fraction of particles that could be considered perfectly oriented at any rate of shear, was essentially the same for intact phage particles and for purified ghosts. The sign of the birefringence was exactly opposite: positive for the ghosts, and negative for the intact phage (Fig. 1). By subtracting these values, and knowing the fraction of particles that could be considered perfectly aligned with the flow, Gellert concludes that the phage DNA is so organized in the head that 10 to 15% of it appears to be perfectly aligned with the long axis of the phage, if the remainder is considered to be arranged at random. The form birefringence associated with the slightly anisometric cavity in which the DNA resides would account for only a small fraction of this. Thus it appears that there is an internal organization of the phage DNA, and this organization is reflected in the birefringence.

In principle more detailed information can be had by X-ray diffraction studies on oriented phage preparations. Such a study was made by North and Rich (1961). Their X-ray diffraction patterns show a high degree of orientation, and a very strong reflection at 3.4 Å which is characteristic of DNA. In an effort to rule out the possibility that these reflections came from extra-particular DNA, they studied the 24 Å equatorial spacing as a function of relative humidity. Their results show that the fibers containing phage do not show the continuous increase in this side-to-side packing distance as the relative humidity approached 100%. With fibers of purified DNA this spacing increases indefinitely as the fiber dissolves. In this regard, the fibers containing phage are similar to DNA-protamine complexes. At present it is not

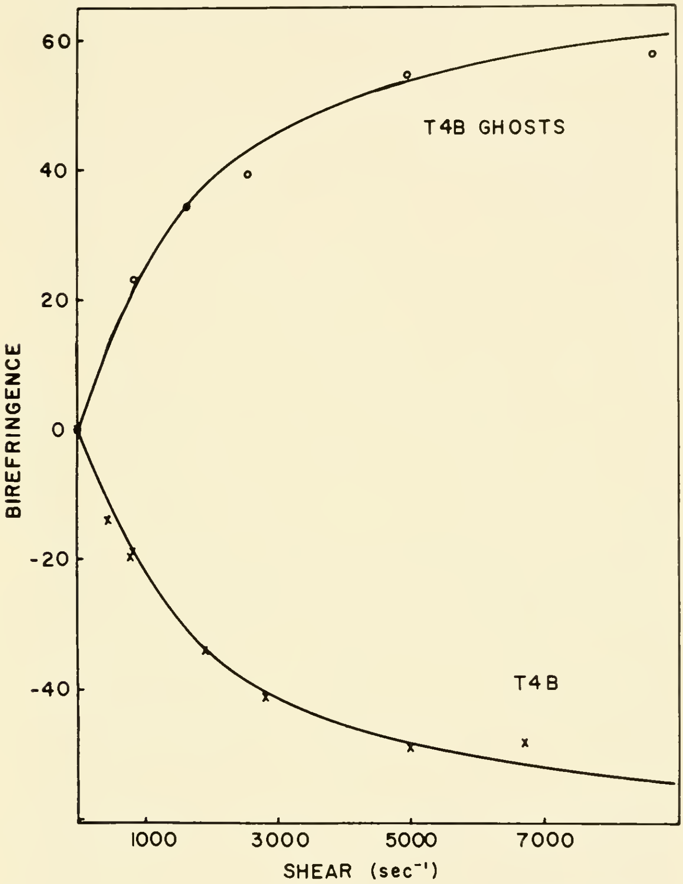


FIG. 1. The birefringence of phage particles and ghosts. This figure shows the way in which the birefringence (mole$^{-1}$ liter$^{-1}$) changes as a function of shear gradient for intact T4 phage (lower curve) and phage ghosts (upper curve). If the DNA were *completely* randomized within the phage head, the birefringence of intact phage would be expected to superimpose approximately with that of the ghosts. (From Gellert, 1962.)

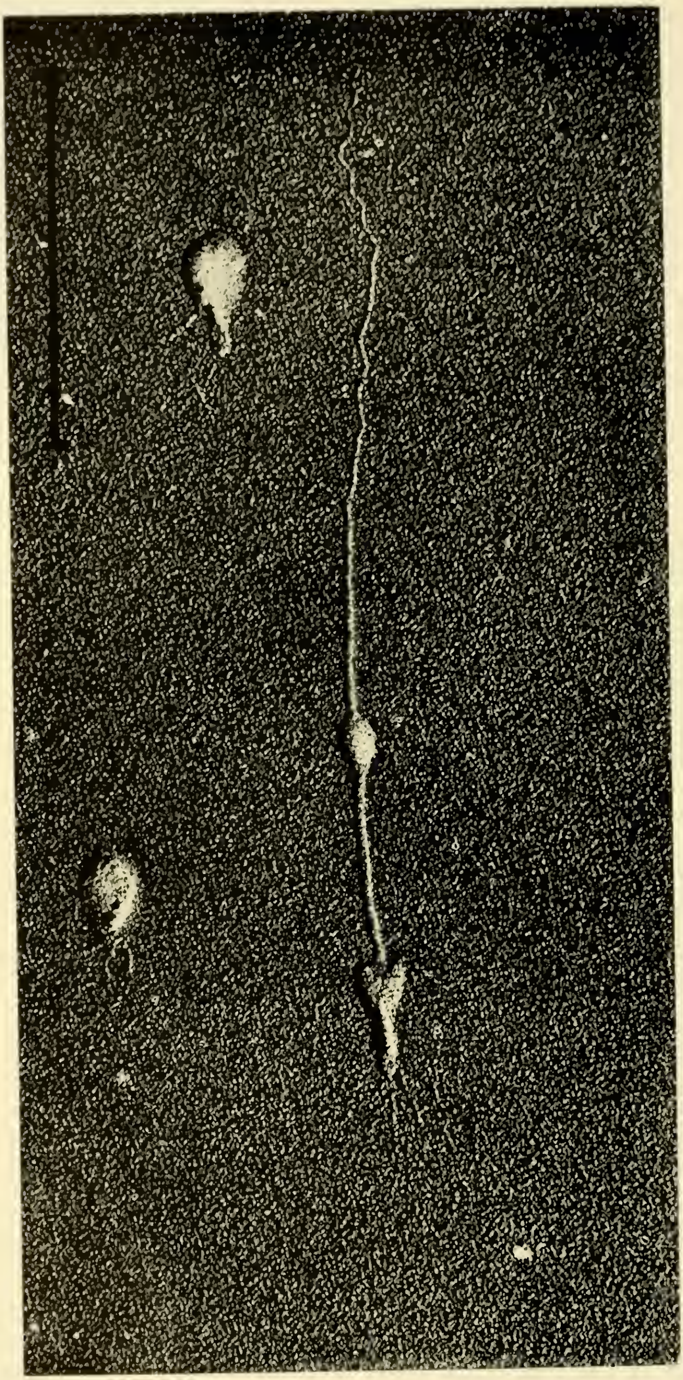


Fig. 2. A T2 phage particle gently disrupted by shear. Many equivalent pictures show the presence of a bulge, which is variable in size and generally about 5000 Å from the broken phage coat. The marker indicates a length of 1 $\mu$. (From Rubenstein, 1960.)

possible to conclude from these studies what proportion of the DNA is oriented parallel to the phage tail.

Similar studies have been made by Maestre and Kilkson (1962), who took great care to avoid and to evaluate DNA ejection. Their results for T2 and M-5 (a long-tailed phage) showed clear 3.36 Å reflections. The M-5 pictures also showed a strong 34 Å meridional spacing and a 52 Å equatorial spacing. This confirms the notion that a significant fraction of the DNA is oriented parallel to the phage tail.

### 2. *Partially Disrupted Phage Particles*

During the course of making electron micrographs of phage particles, occasionally a disrupted particle can be seen with filamentous DNA emerging from it. The first studies of this kind were done by Fraser and Williams (1953). A tangled coil of filaments can be seen very close to the protein ghost from which it presumably emerged. Rubenstein (1960) noticed that phage particles could be broken by hydrodynamic shear on the electron microscope grid in such a way as to draw out the DNA from a tear in the head membrane. This treatment reproducibly resulted in a long DNA fiber, presumably consisting of many DNA molecules, with a characteristic blob or knot of material near the middle of the strand. It is probably safe to conclude that this characteristic knot is a result of some residual tertiary structure of the DNA. But so far no direct information can be drawn about the tertiary structure itself. It is possible that many types of internal organization could result in such fibers with bulges. On the basis of careful measurements of these strands and a consideration of packing requirements, Rubenstein made the suggestion that there exists a tertiary structure based on the formation of cables from the single Watson-Crick duplex strands (Fig. 2).

### 3. *The Electron Microscopy of Phage Maturation*

An important contribution to this problem was made by Kellenberger *et al.*, 1959; Kellenberger, 1962). By sectioning phage-infected bacteria at intervals during intracellular growth, they find that by the 10th minute each cell contains about 5 to 10 dense particles having the characteristic shape of phage particles but somewhat smaller in size. While these dense particles do not form in the presence of chloramphenicol, they are found in abundance shortly after chloramphenicol is removed. On lysis of the cell, these condensed particles are *not* found as intact phage, nor are there enough "empty phage heads" to account for their number. One cannot be completely sure that these condensed units are truly "membraneless" for there is some evidence that the membranes may be broken

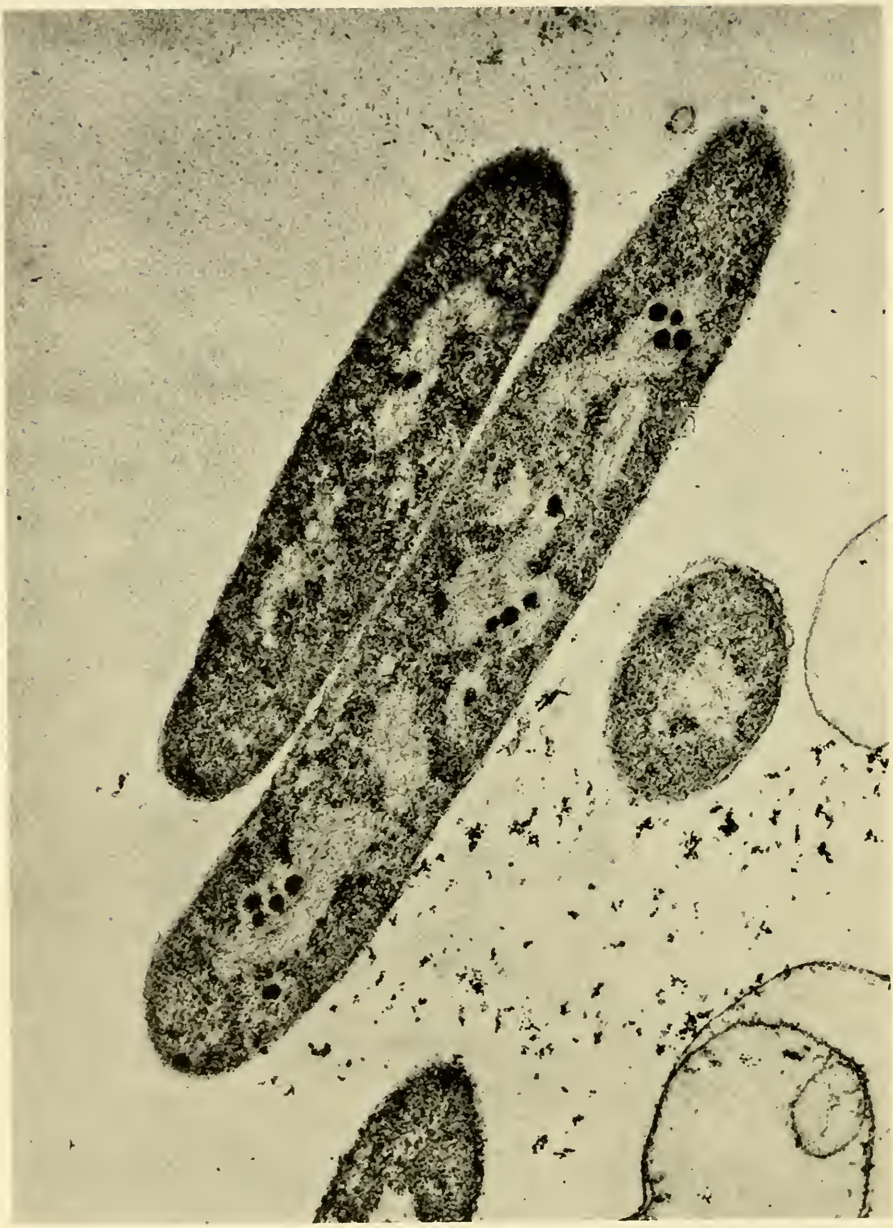


Fig. 3. Organized or condensed phage DNA. The dense areas within the cell are thought to be phage DNA which has been folded and organized prior to the assembly of the head membrane around it. (From Kellenberger *et al.*, 1959.)

down during lysis (Koch and Hershey, 1959). Kellenberger concludes that these "membraneless dense particles" represent DNA which has been organized into a compact form as the result of the action of a still unknown "organization principle" (Fig. 3). To what extent this ability to form a folded configuration is a property of the DNA molecule, and to what extent it relies on the presence of other molecules such as the polyamines (Ames *et al.*, 1958) or the internal proteins (Levine *et al.*, 1958) is a matter of conjecture, as is the role of the head membrane, if present. It seems certain that the internal protein does not bind to free DNA in the presence of 0.1 *M* NaCl, nor is it transferred to progeny phage (Minagawa, 1961). It is of some importance to note that proflavine, an acridine dye which is known to alter the structure of free DNA (Lerman, 1961), also interferes with the intracellular organization of the compact form (Kellenberger *et al.*, 1959).

### 4. *The Question of Single-Chain Regions*

The thorough studies of Sinsheimer (1959a,b) have led to the conclusion that the small bacteriophage, $\phi$X174, contains a single molecule of DNA which is *not* in the form of the Watson-Crick duplex. This phage is about 10 times more sensitive to the decay of incorporated $P^{32}$ atoms, 10 times more sensitive to X-rays, and 25 times more sensitive to UV than T2 phage (after correcting for the 77-fold difference in DNA content) (Tessman *et al.*, 1957; Tessman, 1959). This increased radiosensitivity is in accord with the hypothesis (Stent and Fuerst, 1960) that the lethal event in phage is the complete scission of the DNA molecule; $\phi$X, having only a single chain, would be expected to be broken by *every* $P^{32}$ disintegration. If increased radiosensitivity, and in particular sensitivity to $P^{32}$ decay, reflects single-chain construction of the viral DNA, then these kinds of experiments could yield useful information on the DNA in various phages. The only other phage known to have a high sensitivity to incorporated $P^{32}$ decay (also X-rays and UV) is phage $\alpha$ studied by Graziosi and his collaborators (Celano *et al.*, 1960). However, the DNA excised from this phage has all of the characteristic properties of a duplex molecule such as that taken from T2 or T4 (Cordes *et al.*, 1961; Aurisicchio *et al.*, 1960). In this connection it is important to note that the simple interpretation of $P^{32}$ damage as a "direct" effect needs revision in view of the fact that the efficiency of killing can be decreased substantially by the addition of a free radical trapping agent (Matheson and Thomas, 1960).

A promising approach to the problem of detecting single-chained DNA emerges from the studies of Mayor and Hill (1961). These authors demonstrate that $\phi$X phage particles stain bright red with acridine orange, while T2 phage stain yellow-green (Anderson *et al.*, 1959).

A very sensitive method for the detection of single-chain regions has been perfected by Levine, Murakami, Van Vunakis, and Grossman (1960), based on the complement fixation that attends the reaction of an antibody to DNA which is specific to denatured DNA.

In conclusion, then, one can say that the molecules liberated from T2 and T4 have very little or no single-chain regions while the majority of $\phi$X DNA is single-chained. Molecules of intermediate type have not yet been identified. At present it is not known whether some hydrogen bonds are broken during the folding of the DNA into its compact form.

### 5. *Speculations on DNA Organization*

If the DNA is uniquely folded into a compact structure prior to the assembly of the head membrane about it, as is suggested by the work of Kellenberger, it is reasonable to suspect that at least part of the ability to form specific folds is due to some structural features along the DNA molecule. The presence of such regions along the DNA molecule which are concerned only with the determination of the tertiary folding have been suggested by Mahler and Fraser (1961). From the chemical point of view, there is the unexplained observations of Dunn and Smith (1959) who find about 1 mole of 6-methylaminopurine for every 200 moles of adenine in T2 phage. This corresponds to about 500–600 of such strange bases per phage particle. Since such residues could not form the normal hydrogen bonds with thymine, these bases, if present, might cause an alteration in the normal duplex structure of the molecule. It is interesting to note that the average number of internal protein molecules per phage is about 300 calculated on the basis of the amount of this protein, about 5% (Levine *et al.*, 1958; Minagawa, 1961), and a molecular weight of 15,000 which is based on methionine content (Rubenstein, 1960). In order to compact the DNA into the small volume of the phage head about 300 to 600 turns or folds must be made. The coincidence in the magnitude of these numbers leads one to suspect that they result from some special requirements of the organization of the DNA in the phage particle.

In experiments on the reassembly of protein subunits around TMV-RNA, it seems clear that the RNA plays an important role in the reassembly process. First, because the reconstitution process can occur at higher pH values than will permit the reassembly of the protein units alone, and second, because the length of the TMV rod assumes a unique value—presumably corresponding to the length of the coiled RNA molecule (Fraenkel-Conrat, 1962; Fraenkel-Conrat and Singer, 1959;

Fraenkel-Conrat and Williams, 1955). This is as might be expected in view of the fact that the polyribonucleotide chain is completely surrounded by the protein subunits in the completed TMV particle (Klug and Caspar, 1960).

This experience with TMV leads one to inquire whether the organization of the phage head is strongly dependent on the enclosed nucleic acid. The globular form and the large amount of DNA to be enclosed suggests that not all of the DNA can be in contact with the protein subunits of the head membrane. The existence of "density mutants" in phage lambda indicate that this phage can exist in a number of different forms differing in DNA content (Kellenberger *et al.*, 1961a,b). A variable DNA content seems to be displayed by the different strains of λ-dg transducing phage (Weigle, Meselson, and Paigen, 1959). Likewise the existence of deletion mutations in T4, which seem to be missing a segment of the linkage map, suggests a variability in DNA content (Nomura and Benzer, 1961). One would expect that this difference in DNA content would result in a significant perturbation in the packing arrangement of the DNA. The fact that viable phage particles can be constructed in spite of this alteration in DNA structure suggests that there is some degree of structural independence between the compact nucleic acid and the protein membrane.

In conclusion, it seems clear that the DNA inside the head of bacteriophage has a special organization, but the requirements for this organization are not overly strict because phage particles can be constructed with slightly variable amounts of DNA. The studies of Kellenberger would indicate that the folding and "condensing" of the DNA *precedes* the assembly of the head membrane around the packet of nucleic acid. As to the details of the folding and how folding allows for rapid injection, we remain ignorant.

## B. THE STRUCTURE OF THE LIBERATED PHAGE DNA MOLECULE

### 1. *The DNA from φX174*

The first successful isolation and characterization of the DNA molecule liberated from any bacteriophage was accomplished by Sinsheimer (1959a,b). Measurements of the DNA content by colorometric tests indicated the virus contained 25.5% DNA. Light-scattering molecular weight determinations yielded a value of $6.2 \times 10^6$ daltons for the intact virus. This requires the DNA content to be $1.6 \times 10^6$ daltons. The DNA was extracted and purified by phenol extraction and the molecular weight repeatedly measured by light scattering under various conditions, giving values between 1.6 and $1.8 \times 10^6$. By assuming the absence of any significant systematic error in the light-scattering molecular weight determination, Sinsheimer concluded that there was a single molecule of DNA in each virus particle. A variety of chemical and physical experiments indicates that this DNA molecule does not have the characteristic secondary structure of DNA's from other phages. A large fraction of the amino groups is free for reaction with formaldehyde, and the extinction coefficient does not show a characteristic sharp increase with temperature that is characteristic of DNA's having the duplex structure. It would appear that the molecule consists of a single polynucleotide chain. This molecule is resistant to the single-chain specific exonuclease from *Escherichia coli* (Lehman, 1960) and therefore has no free 3′ end—a finding which suggests a circular form (Fiers and Sinsheimer, 1962).

### 2. *The Autoradiography of T2 and T4 DNA*

Turning now to the T phage, the first information that one would like is the number and size of the DNA molecules making up the phage genome. The molecular weight of the DNA in the even phages is so large that the standard methods of molecular weight determination cannot be trusted; however, autoradiography has proved useful. When $P^{32}$-labeled DNA molecules are mixed with a sensitive photographic emulsion, it is possible to detect single molecules of DNA because they give rise to a cluster of tracks in the developed emulsion. Each track corresponds to the path of a single $\beta$-ray emitted by a decaying $P^{32}$ atom. The first experiments of this kind on T2 DNA resulted in two populations of stars, the first arising from DNA molecules comprising 36–40% of the phosphorus of the intact phage particles, and the remainder in molecules one-fifth this size or smaller (Levinthal, 1956; Levinthal and Thomas, 1957; Thomas, 1959). The validity of this distribution became questionable as a result of the work of Davison, who demonstrated the extreme sensitivity of phage DNA molecules to fragmentation by hydrodynamic shear (Davison, 1959). This work suggested the possibility that larger molecules might have existed, and that they were broken during the dilution and mixing with the melted emulsion.

Further doubt was cast on the autoradiographic results by the work of Hershey and Burgi (1960). By employing a chromatographic column which was able to fractionate large DNA molecules primarily with respect to molecular weight (Mandell and Hershey, 1960), these authors were able to show that unbroken molecules rechromatographed uniformly overlapping the original material, even though the sample selected for rechromatography was taken from the edge of the elution profile. In short, they were unable to demonstrate any heterogeneity in the original unbroken material. These results were in contrast to those obtained with shear-broken molecules, which could be demonstrated to show a certain variability in size. Thus, it seemed likely that whatever the molecular weight of the unbroken material was, it was uniform and not as indicated by the autoradiography up to that time.

This situation was finally resolved by the use of autoradiography and column fractionation by Rubenstein, Thomas, and Hershey (1961). The general scheme of these experiments is shown in Fig. 4.

Phage were labeled with $P^{32}$, then mixed with a large excess of cold carrier phage. The mixture was extracted with phenol following a procedure known to liberate unbroken molecules. The resulting DNA preparation was loaded onto the chromatographic column and the excess $P^{32}$ from the growth medium was washed out with the loading solution. This avoids pelleting the labeled phage which probably results in some radiation damage to the DNA. The DNA is then eluted and, as seen in Fig. 4, the radioactive DNA and the carrier DNA elute in an overlapping peak, which means that as far as chromatographic properties are concerned, the labeled and unlabeled molecules are the same. A single fraction is then selected for gentle mixing with the nuclear emulsion. The resulting population of star sizes shown in Fig. 4E is the same as the population of star sizes that arise from intact phage particles shown in Fig. 4B. Thus, there is a single molecule of DNA which comprises all of the phosphorus of the phage.

By knowing the specific activity of the $P^{32}$ in the growth medium, it is possible to calculate that the minimum molecular weight of this molecule is $130 \times 10^6$, while the phosphorus content per plaque-forming unit sets an upper limit of $160 \times 10^6$. This conclusion is supported by the results of Davison *et al.* (1961), who showed that droplets containing phage particles or DNA molecules containing equal amounts of $P^{32}$ had an equal number of star-forming units. This molecule is not broken down by heating to 70°C, proteases, or column treatment. The kinetics of breakage (Burgi and Hershey, 1961) indicate that there are not a few preferential weak points along the molecule which are more susceptible to shear.

If the mixed DNA preparation is subjected to stirring before loading onto the column, both the labeled and unlabeled molecules are eluted at lower salt concentrations (Fig. 4D). When a sample of this material is examined in the nuclear emulsion, the star population is seen to have a lower mean value (Fig. 4F). By an examination of a number of different fractions of shear-broken preparations, it was concluded that the molecule breaks successively into approximately halves and then into

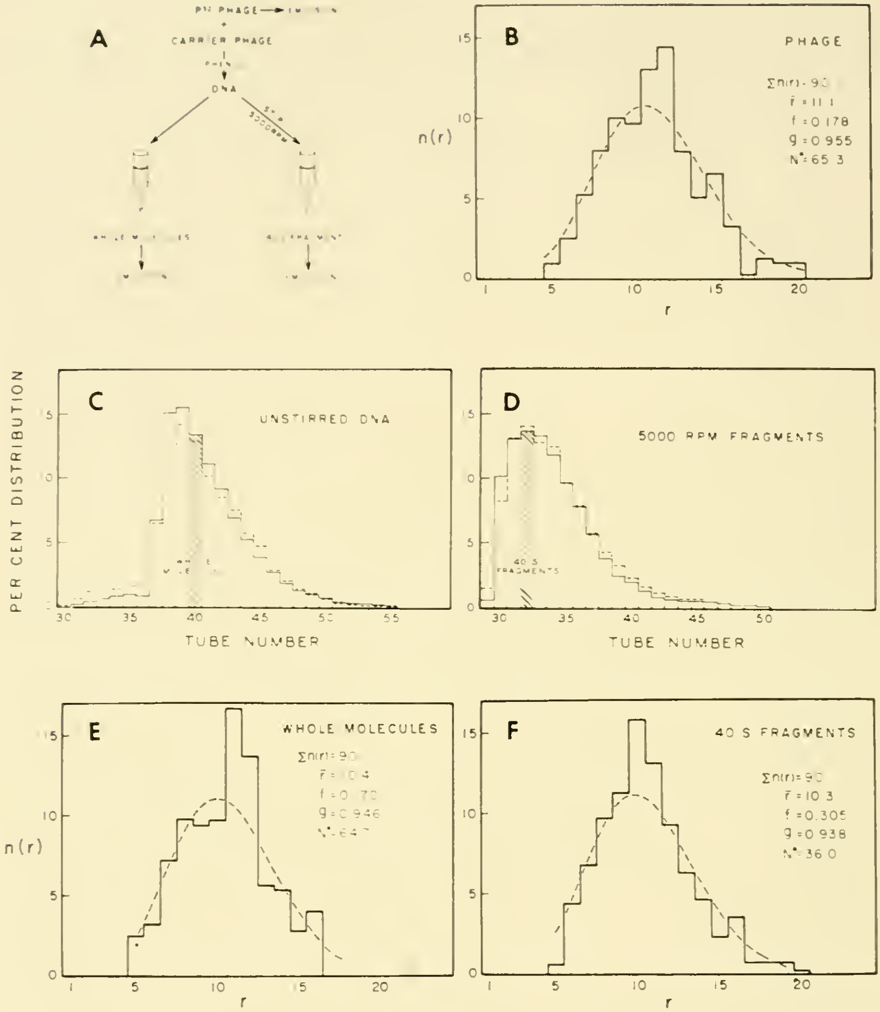


FIG. 4. The results of the autoradiographic experiment. A. The experimental plan. B, E, and F. The frequency distribution of star sizes; the number of stars having $r$ rays plotted against the star size $r$. C and D. DNA chromatograms of intact molecules and broken molecules, respectively. The solid line is optical density; the broken line, $P^{32}$. The shaded fractions were selected for autoradiography. The star population arising from these samples is shown in E and F. (From Rubenstein *et al.*, 1961.)

approximately quarter-molecules when stirred at the critical rates of shear. This is just what one would expect for a linear molecule, while a circular structure would be expected to break repeatedly into fragments which would be substantially smaller than one-half. When a hoopshaped molecule is broken once by shear, its maximum extension would now be twofold greater than in the closed form. This would mean that it would be susceptible to further breakage. The hydrodynamics of this breakage process has been examined by Levinthal and Davison (1961).

The structure of T5 DNA is probably equivalent to that of T2 since the critical stirring speeds that are required to break the T5 molecule are the same as those required to break a T2 DNA fragment of $84 \times 10^6$ daltons (Hershey *et al.*, 1962). In the case of T5, $P^{32}$, and T7 DNA, the sedimentation coefficients are the same as that found for T2 fragments of the same DNA content, if we assume a single molecule per phage particle. However, the sedimentation behavior of these large molecules is not understood, and it would be risky to infer molecular structure from sedimentation coefficients at this time.

The fact that the smallest and largest known bacteriophages have been shown to contain a single molecule of DNA leads one to the obvious hypothesis that *all* bacteriophage contain a single molecule of DNA, and that this unitary structure is a fundamental feature of the function and construction of virus particles.

### 3. *The Mass per Unit Length—the Number of Polynucleotide Chains in the Molecule*

While the autoradiographic results indicate that there is a single linear molecule of DNA liberated from the phage particle, no conclusions can be drawn about the number of polynucleotide chains making up the molecule. Therefore, it would be appropriate to review the evidence that the T2 DNA molecule is a linear duplex of the Watson-Crick type. The X-ray fiber photographs of DNA extracted from T2 and T7 phage show the same characteristic pattern as do DNA's from a wide variety of other sources, except that T2 DNA is unable to form the A form as the relative humidity is reduced; a fact which may be related to the presence of glucose on the hydroxymethylcytosine residues (Hamilton *et al.*, 1959). Thus it would seem safe to conclude that at least a certain fraction of these phage DNA's have the structure derived for the B form of DNA (Langridge *et al.*, 1957, 1960a,b). The equatorial spacing corresponding to the side-to-side packing of the molecules in the fiber increases in a smooth fashion from a minimum value of 20 Å at 50% relative humidity to more than 30 Å as the relative humidity approaches 100% (North and Rich, 1961). This is in accord with expectation for the packing of helical duplex molecules of diameter 20 Å or less and makes unlikely the suggestion that there is a stable association between two duplex molecules, unless one made very special requirements on the packing and on the manner in which hypothetical linkages between molecules changed with changing humidity. The packing requirements of helical duplexes have been examined carefully by Marvin *et al.*, (1961). An analysis of the lattice spacings and the mass per unit length of the rodlike particles has been made by Luzzati *et al.* (1961). Unfortunately, the preparations were mammalian and chicken DNA, not bacteriophage DNA; nevertheless, their results are probably pertinent. They conclude that the particles in solution have the mass per unit length that one would expect for a Watson-Crick duplex, and that this is so under a variety of conditions of concentration and ionic strength. Taking these results together implies that the mass per unit length for T2 DNA would be approximately 210 daltons/Å. For the T2 molecule this would correspond to a structure 62 $\mu$ in contour length.

This conclusion gains strong support from the direct electron microscopic determination of the thickness of the phage DNA molecule by Beer and Zobel (1961). These authors stained the DNA with uranyl acetate, deposited the molecules on a carbon grid with polystyrene latex spheres, and then shadowed with platinum. The molecules were seen to run through the shadow of the spheres. In this region the uranyl-stained molecule can be seen and its thickness measured. Outside the shadow of the sphere the height of the molecule can be measured from the shadow length and known shadow angle. This gives two independent measurements on the same molecule. Both give about 20 ± 5 Å for the diameter of the molecule. Since the entire molecule can be seen in overlapping pictures (Beer, 1961), one would like to know whether *all* of the molecule is in the form of a 20 Å fiber. The vicissitudes of the substrate prevent a confident answer to this question; the best one can say is that there are no *obvious* special features along the limited number of molecules so far seen.

Another way to settle the question of the mass per unit length of the DNA molecule is to measure its total length, and to combine this number with the molecular weight determination.

A direct observation on the total length of T2 DNA molecules has been made by Beer (1961, 1962), who has perfected techniques which allow the molecule to be picked up and stretched out on the electron microscope grid. Using these procedures he has been able to measure 30 different molecules from samples of varying fractional length. When the observed length is divided by the fractional length of the chromatographic fraction, the calculated total lengths range from 30 to 90 $\mu$ with an average of 49 $\mu$. Similar studies on T3 (Bendet *et al.*, 1962) and T7 (Davison and Freifelder, 1961; Beer, 1962) have been made.

Another approach to this problem was made by Cairns (1962), who

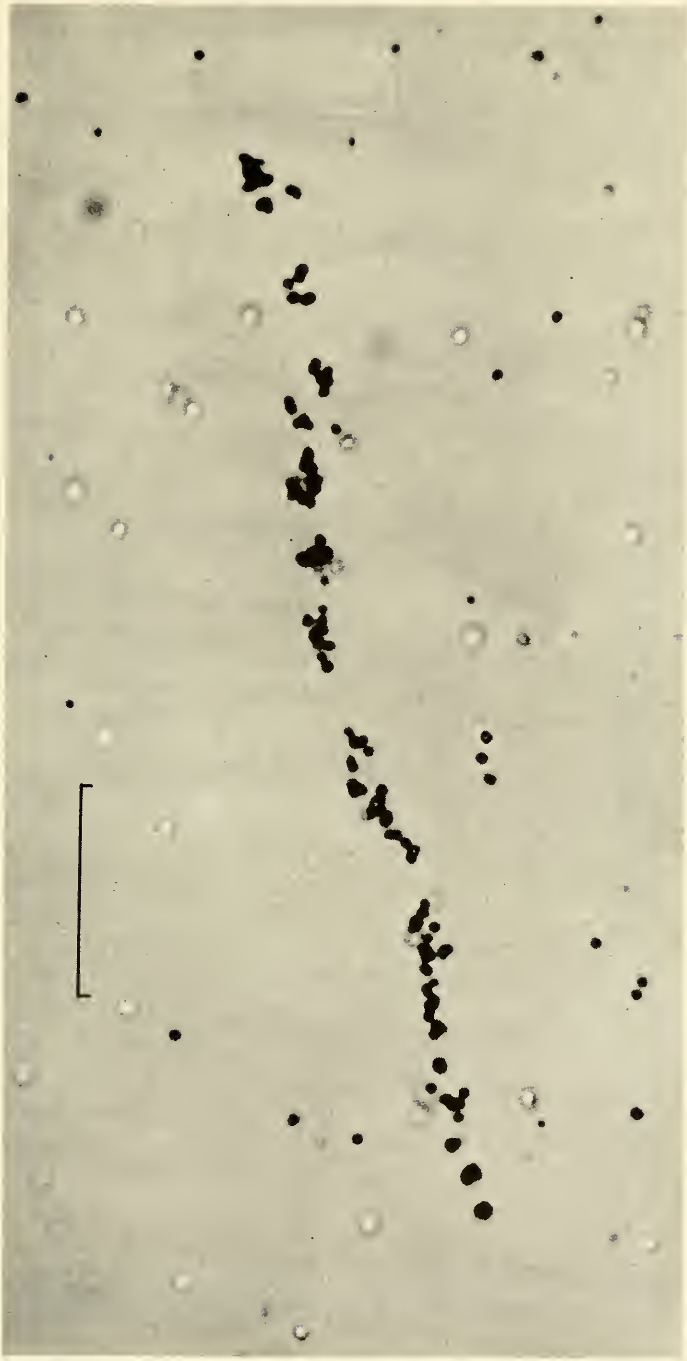


Fig. 5. Tritium autoradiograph of a T2 DNA molecule. The marker indicates a length of 10 $\mu$. (See Cairns, 1962.)

labeled T2 phage with tritium and then mixed with a large excess of unlabeled carrier phage to prevent breakage by shear during extraction. When this DNA is streaked on a glass slide and then covered with stripping film, the localized grains produced by the weak tritium decays are seen to lie in more or less straight lines, the lengths of which are approximately 54 μ (Fig. 5).

Taken together, these studies indicate that the DNA molecule liberated from the phages so far studied have the length and mass that would be expected for a single helical duplex.

### 4. *The Problem of Weak Points or Polynucleotide Chain Interruptions*

Having considered the evidence that T2 and T4 phage DNA is a single linear duplex, one would now like to know if there are any special non-nucleotide residues which unite distinct polynucleotide chains. As mentioned above, stability tests and the kinetics of shear breakage seem to indicate no preferential weak points along the molecule. Therefore, if non-nucleotide linkages exist, they must be at least as strong as the inter-nucleotide linkages themselves. But what about interruptions in the polynucleotide chain; could there be a few breaks located at special or random points in the duplex? Some information on this point has been obtained by sedimentation equilibrium studies done on the unbroken molecules, and the polynucleotide chains produced by the denaturation of the unbroken molecule (Thomas and Berns, 1961; Berns and Thomas, 1961). In these experiments the unbroken molecules were denatured by any of three different methods, in the presence of 1% $CH_2O$. The formaldehyde reacts rapidly with the amino groups exposed by the denaturation, and effectively prevents the reassembly of the polynucleotide chains. This stabilization by reaction with $CH_2O$ is important because the separated polynucleotide chains have a strong tendency to reassociate in solution unless the amino groups are blocked. When a mixture of undenatured whole molecules and denatured whole molecules is banded in CsCl density gradient (in the presence of 1% $CH_2O$), one gets two well-separated, sharp bands. From the variance of these distributions one may calculate a molecular weight value. It turns out that the calculated molecular weight of the polynucleotide chains is just one-half the molecular weight of the undenatured material. This is true irrespective of the mode of denaturation. These results are summarized in Fig. 6. This could only be the case if there were no interruptions in the polynucleotide chains. The method should be able to detect a single interruption if it occurred in the middle third of the molecule. Recent findings indicate that the polynucleotide chains from either end of the molecule have slightly different densities—a fact which would increase the

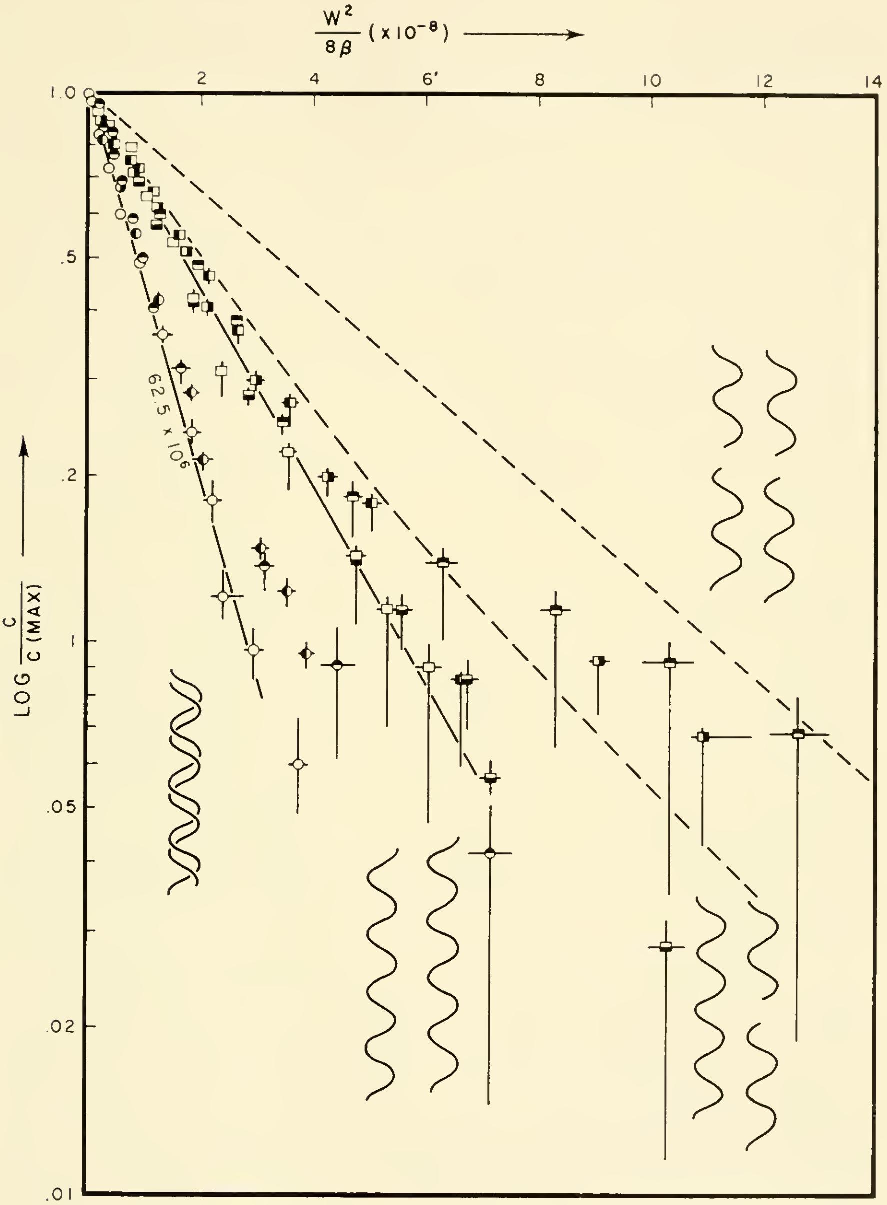


FIG. 6. A comparison of the band forms of T2 and T4 whole molecules with that of formaldehyde-reacted polynucleotide chains derived from them. The Gaussian bands produced by native and denatured whole DNA molecules are shown on the same graph by plotting the log of the band height vs the band width. The whole duplex molecules fall on the steepest curve, while the polynucleotide chains fall on a curve which has a twofold lower slope; a fact which indicates that their relative molecular weight is lower by a factor of 2. (From Berns and Thomas, 1961.)

ability to detect nicks by producing a still broader band (Thomas and Pinkerton, 1962).

Two control experiments should be mentioned. If a few interruptions in the polynucleotide chains are produced by the limited reaction with DNase, then the shorter chains are easily observed by the disproportionately broad band produced by the single polynucleotide chains, whereas the undenatured duplex has about the same band width as before. In order to rule out the possibility that formaldehyde is forming cross-links between polynucleotide chains, $N^{14}$–$N^{15}$ hybrid molecules were produced by reannealing. On subsequent denaturation in the presence of $CH_2O$, the chains of different density were cleanly separable, which meant that the formaldehyde was not causing them to be linked together (Thomas and Berns, 1962).

When the continuity of the molecule is maintained by a single chain, one would expect that its tensile strength would be one-half that of the duplex. Thus, interrupted molecules might be more fragile to shear than uniform duplexes. The major consequence of the decay of an incorporated $P^{32}$ is the destruction of the phosphodiester linkage and thus the interruption of the polynucleotide chain. Therefore, it is possible to produce artificially interrupted molecules in this way. This expectation was confirmed by Levinthal and Davison (1961), who found that T2 molecules which had suffered a substantial decay of $P^{32}$ were broken at lower rates of shear than was required to break undamaged molecules. Since interruptions apparently *do increase* the shear fragility of the molecule, this indicates that there were *no* interruptions in the original molecule in those regions which can be broken by shear. Since the maximum tension is developed in the middle, these experiments suggest that there are no natural interruptions near the middle part of the molecule.

The sedimentation velocity analysis of the chromatographic fractions of the first breakage product of whole molecules (Burgi and Hershey, 1961) shows that there is a single, approximately Gaussian distribution of molecular sizes produced by stirring. This distribution is not altered by prolonged stirring at the critical speed, and would only be produced if the most probable point of breakage were in the center of the molecule, and if the probability of fracture decreased in an approximately normal manner on each side of the middle. This would indicate that the probability of the breakage is governed by the hydrodynamic factors in the region of the stirring blade, rather than weak points located at a special location. The same kind of analysis was applied to the breakage of half-molecules by stirring at a higher (critical) speed. This produced a normal distribution of molecular lengths about one-quarter the original size. On the basis of these arguments, and others based on the critical rates of shear required to break molecules of different sizes, these authors come to the conclusion that there are no weak points located at either special or random locations along the molecule.

### 5. *Conclusions*

The studies on the DNA from T2 and T4 show that there is a single, stable, linear DNA molecule which comprises all of the nucleic acid in the phage particle. Insofar as is known, all bacteriophage may contain a single molecule of nucleic acid. The length of the molecule (in T2) is approximately what one would expect for a duplex molecule. When the polynucleotide chains are separated by denaturation in the presence of formaldehyde, the high molecular weight of the polynucleotide chains indicates the absence of interruptions in the continuity of the chains. Thus, the bacteriophage DNA molecule could be a simple duplex over its entire length.

## C. THE GENETIC STRUCTURE OF PHAGE

A central concept that has developed in microbial genetics is that the genetic structure, as revealed by the compiled data from recombination experiments, is related in an understandable way to the physical structure of the DNA molecule, and that the physical structure is reflected by the processes of synthesis and recombination in the form of a genetic linkage map. The processes of "synthesis" and "recombination" are themselves not understood, and the clarification of these events are among the major objectives of molecular genetics. In this section some features of the genetic structure of phage will be considered in their relation to the physical structure of the phage DNA molecule.

### 1. *The Circular Genetic Map in T2 and T4*

On the basis of a sensitive linkage test, Streisinger and Bruce (1960) have come to the conclusion that all of the known genetic markers in T2 and T4 are linked to form a continuous genetic map. In these experiments three markers were employed, two of which were known (or could be demonstrated) to be linked, and a third the linkage of which was to be tested. By selecting for observation a class of progeny phage which showed recombination between the two markers that were known to be linked, they could then examine these phage for the third character. If this third character were totally unlinked, one would expect to find it associated with the most proximal of the linked pair with a frequency depending solely on the relative multiplicities. The experiment can be done with greater sensitivity by comparing the results of two crosses, the first (Cross I) with the multiplicity of the parent bearing the markers *a*, *b*, and *c* about 10 times that of the minority parent $a^+b^+c^+$. A second cross (Cross II) was performed with exactly the *same*, but *opposite* multiplicities. In Fig. 7 the relationships are depicted.

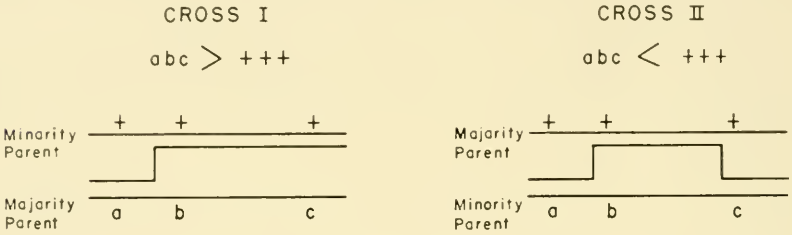


FIG. 7. The types of crosses which give a sensitive test for linkage in phage. The recombinant type indicated should be equal in the two crosses if the *c* marker were unlinked from the linked *a* and *b* markers. The recombinant type shown in Cross I was actually substantially higher than that shown in Cross II. (From Streisinger and Bruce, 1960.)

In this figure the recombinant type is indicated (for convenience only) by a copy-choice scheme. In these two crosses the progeny phage will contain a certain number of phage of the recombinant type indicated. If this number is divided by the total number of a $b^+$ recombinant types, irrespective of whether they contain *c* or $c^+$, the resulting ratio should be identical in both crosses—*if* the *c* marker is truly unlinked. In all experiments that were performed in both T2 and T4, the ratio obtained in Cross I was substantially higher than that obtained from Cross II—a fact which indicated that the *c* marker is linked to the other two. These conclusions are completely free of assumptions concerning either the kinetics of replication or of recombination. Using this technique with various markers thought to be unlinked, it was demonstrated that they were all linked to each other. This unitary genetic map has been confirmed using the many temperature-sensitive mutants which can be found distributed over the entire map. The closer spacing of these markers makes the sensitive test described above unnecessary (Edgar, 1961).

By a continuation of the type of crosses done by Streisinger and Bruce, it was found by Streisinger, Edgar and Harrar (1961) that markers previously thought to be at the extremes of the genetic map were themselves linked to each other. Thus the genetic map in T2 and T4 must be considered to be formally circular. This is in contrast to the *linear* structure of the excised T2 or T4 DNA molecule. At present it is not yet known whether any other smaller virulent phages have a circular map. In the exploration of the genetic structure of any organism one is restricted to that region of the map which contains recognizable markers. The available markers in lambda are in a region which is only about 20 units long, and there is no evidence of circularity.

The discovery of the circular map in T2 and T4 makes it quite certain that additional map segments will not be added as additional markers are discovered. Thus one can state with confidence that the map in T2 is known, unitary, and circular. The DNA molecule liberated from the phage, containing essentially all of the nucleotides in the virus, is unitary, but almost surely linear. Thus, at present, one would like to know the relationship between the circular map and the linear molecule. Perhaps each DNA molecule has genes in the same order but the beginning and end points are different. Such a circularly permuted molecule would make the genetic results easily understood. Alternatively, the molecule and map could be unique but two (or an even number of) cross-overs *always* occur during a recombinational event (Stahl, 1962). Some of these possibilities can be tested. At the present time, one is limited by the fact that no sensitive transformation assay for T2 segments has been developed.

### 2. *Transformation by Lambda-dg DNA*

In the case of phage lambda the situation is more favorable as a result of the work of Kaiser and Hogness (1960) and Kaiser (1962) who have demonstrated a transformation system with lambda-dg, and lambda DNA. When a bacterial strain which is lysogenic for lambda is induced by ultraviolet light, the vegetative phase of growth is initiated and new phage particles are synthesized and subsequently released. Approximately one in $10^6$ of these new particles is lambda-dg. These phage are different from ordinary lambda in a number of ways. This phage particle can transduce the galactose genes from the bacteria in which it was grown to the bacteria it infects. It is defective in the sense that it is missing a sizable fraction of its genetic map (Arber, 1958) and it has a very small chance of establishing the lysogenic condition compared wth normal lambda. Independently arising lambda-dg phage can be shown to differ in their density when banded in CsCl, indicating that they have a variable DNA content (Weigle, Meselson, and Paigen, 1959). It is currently thought that these lambda-dg phage are the result of a substitution of bacterial genes for some of the phage genes. This exchange appears to be confined to a special region of the lambda map, and appears to be unequal in the sense that the resulting lambda-dg may have somewhat more (or somewhat less) total DNA than ordinary lambda.

Kaiser and Hogness purified the transducing phage by making use of a property of lambda-dg, namely, that it has a different density than ordinary lambda. The DNA was then extracted from these phage by the phenol technique, and the resulting DNA could be shown to produce Gal$^+$ colonies when added to appropriately lysogenized Gal$^-$ cells in the presence of "helper" phage. While the details of this experiment are somewhat technical, the net result is a *transformation* of the recipient cell to Gal$^+$ which is scored on EMB-galactose plates. A further examination of these cells has revealed that they are Gal$^+$ because they have become lysogenic for lambda-dg. If this is true, these transformants should also contain other genes (outside the dg region) as well. This was tested by looking for the cotransfer of *c* and *mi* genes which are located at some distance from the end of the dg region. It was found that bacteria which had received the Gal$^+$ also had a high probability of receiving *both* the *c* and *mi* genes as well. This suggested strongly that the molecule which was responsible for the transformation contained the *entire* lambda-dg chromosome.

### 3. *Transformation with DNA from Ordinary Lambda*

These studies on lambda DNA were continued by Kaiser (1962), who developed an assay for ordinary lambda DNA; in this case the assay was based on the production of infectious phage particles rather than the transformation of a recipient cell. Again the role of "helper" phage is important. A non-lysogenic strain of *E. coli* K12 was infected with helper lambda (bearing the $i^{434}$ marker) at a multiplicity of 10. After a suitable period for adsorption, the phage-infected bacteria were washed and then mixed with the diluted DNA solution. This DNA was prepared by phenol extraction of a purified preparation of lambda phage bearing the appropriate genetic markers. After a period of approximately 40 minutes DNase was added to digest any extracellular DNA and appropriate dilutions of this mixture were plated on indicator bacteria to characterize the progeny phage genetically. These indicator strains will not record phage which are produced by the "helper" phage, but they will record phage bearing one or more genetic markers furnished by the DNA. The map of the lambda phage from which the DNA was prepared is given in Fig. 8. It was found that when the DNA contributed genes, it contributed *all* of the markers on the above map. There was the apparent simultaneous transformation of all of the markers, and in this regard it was similar to phage infection. Thus the infecting DNA molecule comprises the entire chromosome of the phage. When this DNA preparation is stirred, the molecules are broken near the middle, causing them to sediment more slowly in a sucrose density gradient. The ratio of the sedimentation constant estimated before and after stirring indicates that the molecular weight is lower by a factor of 2. The biological assay of the stirred DNA now shows that most of the phage progeny contain the $i^\lambda$ marker, but very few now contain the *h* marker; there is no longer the simultaneous transfer of *all* markers. By replating the resulting phage, it is possible to examine all the markers shown in the map in Fig. 8. It turns out that both the *s* and *mi* markers are transferred with essentially the same frequency as is the *i* gene. Thus it appears that stirring, a treatment which breaks the molecule near its mid-point, actually causes a separation of the map between the *h* and *s* markers. This would be expected if the map sequence and the nucleotide sequence were equivalent. The fact that both the map and the molecule are broken near their midpoints suggests that the available markers are distributed over most of the length of the molecule.

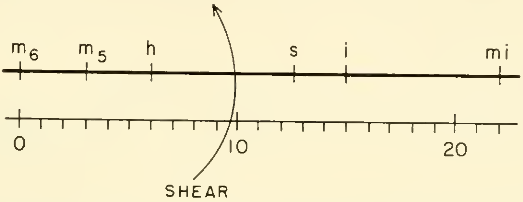


FIG. 8. The map of lambda. The result of hydrodynamic shear is to eliminate the simultaneous transformation of both *h* and *i* genes. The simplest interpretation of these experiments is depicted by the arrow labeled "shear."

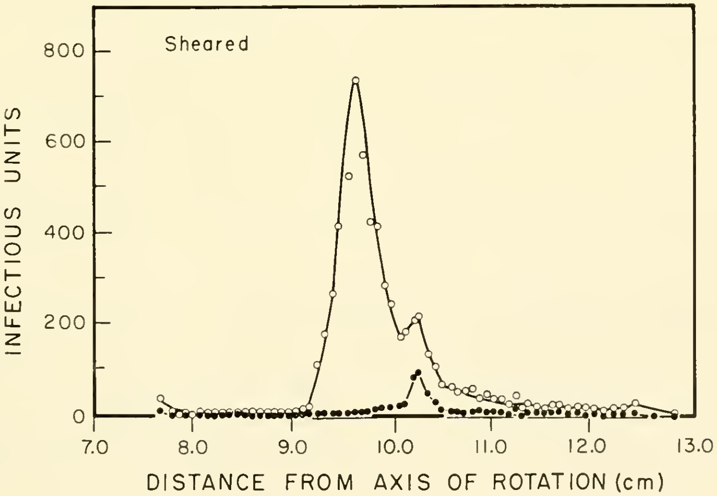


FIG. 9. The separation of DNA fragments bearing the $i^\lambda$ gene (○) from unbroken molecules bearing the pair $hi^\lambda$ (●) by sucrose gradient centrifugation. (From Kaiser, 1962.)

By employing the sucrose gradient sedimentation technique, it was actually possible to physically separate the unbroken molecules bearing the *h and* $i^{\lambda}$ genes from the broken molecules which can transform the $i^{\lambda}$ gene *only*. This is shown in Fig. 9.

If extensions of this kind of experiment can be made to the other genes in lambda, and to other genetic systems, the results may prove what everyone has hoped: that the map sequence and the nucleotide sequence is the same, and that genetic structure is a simple manifestation of physical structure.

### 4. *The Structure of Heterozygous Phage Particles*

When a bacterial cell is infected with two bacteriophage T2, one of which is a mutant and the other wild type, approximately 2% of the resulting progeny will contain both alleles. These *heterozygote* particles were discovered and analyzed by Hershey and Chase (1951). In general these heterozygotes (HETS) will give rise to *both* types of phage on further growth, and this is the way in which they are detected. So far as is known it is possible to find particles which are heterozygous for any marker tested, although the *r* (rapid lysing) marker is very convenient because HETS will produce plaques containing both *r* and $r^+$ phage. These plaques have a characteristic mottled appearance. Pairs of markers which are closely linked are often heterozygous for both markers simultaneously, while an unlinked marker shows only a slightly higher than usual (2%) chance of also being heterozygous. Further work showed that phage which were heterozygous for a central marker were in general recombinant for markers on either side of the heterozygous region (Levinthal, 1954; Trautner, 1958), a fact which indicated that these structures could play an important part in the formation of recombinants. These findings were extended by Edgar (1957), who was able to demonstrate that there was a higher probability of recombination between closely linked markers when cells were infected with heterozygotes (for these markers) than was the case for mixedly infected bacteria. Perhaps the most important of Edgar's finding was that the heterozygote particle gave rise to a *distribution of recombinants* which suggested that the heterozygote *persisted* or *multipled* within the infected bacteria in such a way as to produce recombinant particles at any time during the latent period. Thus heterozygotes were pictured as a more permanent feature of the phage DNA, and not as structures which were *obliged* to segregate on replication. These findings led to the investigation of the fine structure of the heterozygous region by Doermann and Boehner (1962), who developed techniques to test for six different closely linked *r* (*rII*) mutations and a distant *tu* marker.

Phage bearing 6 markers were crossed with wild type, and the mottled plaques were resuspended and replated. The *r*-type plaques were then tested to learn exactly which of the 6 *r*'s are represented on the plate, for these *r*'s must have been originally found in the heterozygous phage. They found that these heterozygous particles must have contained all 6 $r^+$ markers (otherwise they would not have been detected as mottled plaques), and in general 1, 2, 3, 4, 5, or 6 *r*'s usually in contiguous groups. For example, if *a*, *b*, and *c* were present and *d*, *e*, and *f*, were absent it was concluded that the heterozygous segment must have terminated in the region between *c* and *d*. By further examination of these heterozygotes which had a single termination in the region covered by the 6 *r* markers, Doermann and Boehner demonstrated that *r* markers near the termination of a string of *r*'s were not present in the mottled plaques as frequently as markers near the extremes of the *r* region. This would only be expected if the existence of a termination somehow caused the markers near the termination to be copied less frequently (in preference to the wild type alleles). From these experiments Doermann and Boehner conclude that the physical structure of these heterozygotes is probably not a simple duplex DNA molecule, but rather a more complicated structure which actually has *more* DNA than a non-heterozygous particle. Their proposal is shown in Fig. 10.

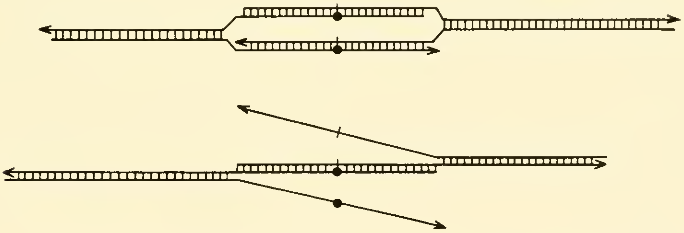


FIG. 10. Proposed structures for heterozygous DNA molecules. (From Doermann and Boehner, 1962.)

This model receives some support from the findings of Doermann (1961). When progeny phage containing recognizable heterozygotes are banded in CsCl and fractions taken, the heterozygous particles are 15 times as numerous in the denser edge of the band, indicating that they could indeed contain more DNA than the average phage particle, or that there is some other special reason for these particles to be denser.

The termination of the genetic strand in the heterozygous region and the finding that HETS have an increased density in CsCl suggests that this discontinuity should be recognizable by physicochemical experiments. So far, the banding studies of Berns and Thomas (1961) indicate that the polynucleotide chains are continuous and noninterrupted. However, their conclusions refer to the *majority* of the molecules. Moreover, their technique would not detect interruptions near the ends of the molecule. Finally, an overlap structure of the kind shown in Fig. 10 could result from the closure of a linear molecule on itself to form a circular structure, a hypothesis which would allow only a single heterozygous region of this form per molecule. Therefore, at this point it is not clear whether there is a conflict between the genetic and physical experiments or not.

### 5. *Conclusions*

At the present time one may picture the phage DNA molecule as comprising all of the DNA in the phage particle. After release the T2 molecule is found to be linear and over 50 $\mu$ in length. The genes in lambda appear to be arranged on a linear molecule and can be separated by breaking the molecule by shear; one expects (but as yet no evidence) that the genes in T2 are in map sequence along the T2 molecule. There are two aspects of the genetic structure of T2 which as yet have no physical interpretation: (*1*) the formal circularity of the T2 linkage map, and (*2*) the existence of a partly stable heterozygous region which displays genetic discontinuities. It is important to note that the combined length and mass determinations of the T2 DNA molecule only allow one to conclude that the majority of the molecule is duplex. Direct visualization in the electron microscope supports this conclusion. It is still possible that as much as 10% of the molecule may have a more special or elaborate structure which may perhaps correspond to the physical heterozygous region, although it must be said with equal weight that there is at the moment no physical evidence in favor of these regions.

## D. THE DISTRIBUTION OF PARENTAL PHAGE DNA TO PROGENY PHAGE

With the finding that the DNA in mitotic chromosomes is in the form of two subunits which, upon replication, separate and become associated with newly synthesized subunits (Taylor *et al.*, 1957; Meselson and Stahl, 1958a,b), attention turned quite naturally to the question of the physical and chemical structure of these subunits. In view of the small amount of DNA in bacteriophage, it seemed likely that phage chromosomes were of a relatively simple structure. This has to a large extent proved so, and has been the subject of previous sections of this chapter. In this section the manner in which the DNA from parental phage segregates and is distributed into progeny phage molecules will be considered. This distribution of parental DNA should be related to the structure of the phage molecule and its mode of replication.

### 1. *Transfer of φx DNA*

This virus shows no or very little transfer of parental DNA to progeny phage (Sinsheimer, 1961).

### 2. *Transfer of T7 DNA*

When coli are infected with T7 phage labeled with $N^{15}$ and $C^{13}$, the parental DNA that is labeled with these heavy isotopes is transferred to progeny phage particles. When the progeny phage are examined by density gradient centrifugation, there appears a band corresponding to phage with 24% of their atoms substituted by heavy isotopes (Meselson, 1960). Since T7 phage are about 50% DNA (Putnam, 1953; Davison, 1962) this means that some phage particles are formed which have half newly synthesized DNA and half parental DNA. The total DNA content of this phage is approximately $20 \times 10^6$ daltons (Davison and Freifelder, 1962) and there is evidence to support the notion that it is in the form of a single molecule (Thomas and Pinkerton, 1962). *If* the biological formation of hybrid T7 phage could be accomplished it could now be established that these phage DNA molecules which have hybrid density are simple duplexes of polynucleotide chains, and if it could be shown that the heavy and light subunits cannot be separated by breaking the molecule by shear, but that separation of the density label *does* occur by denaturation and separation of the polynucleotide chains, then it would be permissible to conclude that the transferred subunits are single polynucleotide chains, and that they are associated with a newly synthesized polynucleotide chain to form a *hybrid duplex*.

### 3. *Transfer of T2 or T4 DNA*

In contrast to the results of T7 the parental DNA of a T2 or a T4 phage is broken into smaller pieces and incorporated into the DNA molecules of *many* progeny phage. The transfer is dispersive. The early work on this subject has been reviewed by Sinsheimer (1960). These experiments by Levinthal (1956), Hershey and Burgi (1956), and Stent, Sato, and Jerne (1959) presented evidence that a sizable fraction (40 to 60%) of the parental DNA was transferred to a relatively few progeny phage particles in units which were as large as 15 to 20% of the parental phage. The remaining parental DNA appeared to be even more highly dispersed to a large number of phages in pieces no larger than 0.3%. This transfer of apparently very large pieces of DNA suggested that at least *part* of the phage chromosome was replicating semi-conservatively.

This view was supported by the findings of both Stent *et al.*, and Levinthal that these relatively large pieces were not further dispersed on transfer to a *second* generation of progeny phage.

With the advent of density labeling techniques, two groups began to reinvestigate these problems: Kozinski, using 5-bromodeoxyuridine (BU), and Roller and Meselson employing $N^{15}$ and $C^{13}$ (Kozinski, 1961; Roller, 1961). In the case of Kozinski's experiments, normal $P^{32}$-labeled phage were allowed to singly infect "heavy" bacteria growing in a medium containing 5-bromodeoxyuridine. Progeny phage were collected and DNA extracted by the phenol procedure. This DNA was banded in CsCl and the $P^{32}$ (which came from the parental DNA) was now found to be in a band which overlapped the band for totally new or "heavy" DNA. This experiment can be done in reverse as a result of perfecting methods to grow highly infective, uniformly BU-labeled phage (Kozinski, 1962). In Roller's experiment "heavy" phage are allowed to infect "light" bacteria and the progeny DNA molecules are examined by centrifugation in CsCl. Again the conclusion is the same, namely, that all of the parental label is now found in molecules having a density which is essentially the same as that of the newly synthesized "light" DNA. Although it is difficult to estimate the upper limit to the size of the transferred pieces, it seems clear that only a small fraction of the parental DNA now resides in progeny molecules where it makes up as much as 10% of the total molecule. Following breakage by shear (Roller, 1961) or sonication (Kozinski, 1961) the parental label now assumes a position in the band corresponding to molecules made up of half-heavy and half-light DNA. When the majority of the DNA has been sheared to a molecular weight of about $10 \times 10^6$ (as estimated by sedimentation coefficient), the majority of the parental DNA is found at the half-heavy position. One would like to conclude directly that these fragments are hybrid duplexes, in view of the fact that the majority of the phage molecule is duplex. However, the possibility still exists that the parental subunits now preferentially reside in some special regions of the molecule, the structure of which is not known, and the existence of which cannot completely be ruled out at this time.

Unfortunately, there are no published experiments in which the density labeling techniques have been employed to examine a second generation of phage. There is at present very little information available from these experiments on the *distribution* of the sizes of the transferred pieces, and even less on the perpetuation of this distribution during another cycle of infection.

There is a conflict between the earlier results and those derived from the density labeling experiments just mentioned. It could be one of degree, however, for all experiments indicate that the transfer is dispersive, and that the difference is in the degree of dispersion. It is not unlikely that the degree of dispersion depends on physiology of the infected bacteria. At this point there has been only limited effort to reconcile these apparent differences. The transfer studies using $P^{32}$ autoradiographs have recently been reinvestigated by Kahn (1961), who finds in general smaller and more heterogeneous parental pieces in the transferred phage than was found by Levinthal.

### 4. *Transfer of Lambda DNA*

At the present time there have been no published experiments which have attempted to trace the distribution of parental DNA into progeny lambda DNA molecules, although density gradient studies have been done on whole phage. Meselson and Weigle (1961) have shown that the progeny lambda phage collected from a single cycle of infection by an $N^{15}$–$C^{13}$-labeled parent produce *some* infectious phage which have a density corresponding to phage particles made of totally new protein and totally heavy, parental, DNA ("conserved" phages), and another peak having the density expected for phage having half parental and half newly made DNA enclosed in newly made protein ("semi-conserved" phages). Phages with recombinant genotype were shown to contain parental DNA because they were found at certain positions between those expected for fully conserved and totally light or newly synthesized phage. Experiments of a similar nature have been performed by Kellenberger, Zichichi, and Weigle (1961a) using phage which have natural density difference as a result of a difference in DNA content (Kellenberger, Zichichi, and Weigle, 1961b). Again phages of recombinant type would be shown to contain preferred quantities of parental DNA. These experiments have been discussed more fully in Chapter II. It is important to note here, however, that the transfer in lambda is quantitatively different from that found in T2 or T4 phage. The parental DNA is not dispersed into pieces which are small (compared to the total DNA content of the lambda phage) and dispersed into many progeny phage.

## II. BACTERIAL DNA

At the present time we have much less precise knowledge of the organization of bacterial DNA's than we do of bacteriophage DNA's. Again what information is available has come from the interpretation of cytological studies, growth and transfer experiments, the study of purified DNA, and genetic experiments.

### A. CYTOLOGY OF BACTERIA

The excellent review by Kellenberger (1960) has covered this subject in detail, and only some major conclusions will be considered here. He concludes that in bacteria the DNA is organized into well-defined areas called nuclear vacuoles which are variable in shape and do not appear to have a membrane. These vacuoles are filled with a fibrillar material about 20–60 Å in diameter; the dense chromosomes which are sometimes seen are thought to be a coagulation artifact. As yet, electron micrographs are not able to reveal anything about the number of DNA molecules in the fibrils, or anything about the linkages between DNA molecules, although some recent work in this area appears promising (Kleinschmidt, Gehatia, and Zahn, 1960).

### B. GROWTH AND TRANSFER EXPERIMENTS

Following the experiments of Meselson and Stahl (1958a,b) which demonstrated that the replication of *E. coli* DNA took place by the formation of a structural unit which was one-half old and one-half newly synthesized material, subsequent work has been aimed at *(1)* discerning the *largest* structure which will replicate by this semi-conservative fashion, and *(2)* identifying precisely the structure of the DNA molecule which contains equal amounts of old and new material. The first of these topics has been investigated by Forro and Wertheimer (1960) by tritium autoradiography. His experiments have been considered in Chapter II. For the purposes of this discussion the salient point is that by using young, small cells of *E. coli* 15T$^-$, the DNA seems to be synthesized as though the bacteria contained only *two* DNA-containing structures which can survive many rounds of replication. These findings have been verified by Van Tubergen and Setlow (1961) using tritium autoradiography. Exponentially growing cultures of coli which probably have two nuclear vacuoles were found to have *four* structures which were conserved during further replication. This would be just as expected if each nuclear vacuole contained a single long DNA molecule, or a molecule punctuated by a special non-DNA unit ("linkers") at various points (Kellenberger, 1960; Freese, 1958). Thus, while the details of the structure of such a unit are still unknown, one may consider the "nuclear vacuole" as seen in electron micrographs to consist of two DNA-containing units which may (but rigorous proof is still lacking) be the two polynucleotide chains of a very long DNA molecule.

Further evidence on this point has come from the study of an *E. coli* which simultaneously requires thymine, arginine, and uracil (Maaløe and Hanawalt, 1961; Hanawalt *et al.*, 1961). The starting point for these investigations is the thymineless death phenomena discovered by Cohen and Barner (1954). When thymine is abruptly removed from a mutant strain of coli ($15T^-$) which has a requirement for thymine, the cells gradually lose their viability in an exponential fashion for over four decades. In the case of the triple mutant $T^-A^-U^-$ thymineless death proceeds essentially as before when the cells are incubated in a medium lacking T, A, and U *except* that there is a residual fraction (3% to 10%) of the cells which is immune to thymineless death. This fraction cen be *increased* to 100% of the cells by a 90-minute preincubation in a medium containing T, but lacking A and U. These authors make the plausible interpretation that this preincubation period allows the cells which have begun DNA synthesis to *complete* the templating of a genome. Before another round of templating can begin, there are some essential steps which require *arginine* and uracil, which were missing from this medium. This suggests that thymineless death is the result of *attempted* DNA synthesis which can take place in the absence of thymine only at a very slow rate and with a high incidence of errors, some of which are lethal.

If it were true that only a completion of a round of replication were permissible in the presence of thymine, and in the absence of arginine and uracil, then the resulting molecules after being broken during extraction should have hybrid density if 5-bromouracil were added in place of thymine. This hypothesis would predict that no fully heavy molecules would be found even on prolonged incubation with 5-bromouracil. On banding the extracted DNA, two distinct, symmetrical bands were found which corresponded to the density of normal DNA and to DNA of hybrid density. Thus, segments of the expected form were found. This conclusion is weakened somewhat by the control experiments which indicated that it was difficult to form uniformly BU-labeled DNA; a fact which is probably related to an inhibitory effect of bromouracil (Maaløe, 1962).

Taking these results, together with those of Forro, strongly suggests that the bacterial chromosome is a single structure of fixed DNA content, and that the two independently labeled units that emerge from the studies of Forro are united in equal proportions in such a way as to produce hybrid material in the density gradient. Since the hybrid DNA can be broken extensively by sonication and yet retain its hybrid density, one is forced to admit that traverse breaks do not separate the subunits (Rolfe, 1962). This is exactly what one would expect if the two independent labeling units were the complementary polynucleotide chains making up the duplex DNA molecule. On the other hand, one would expect the same thing from a pair of duplex molecules united at frequent intervals, one of which was a pair of old polynucleotide chains, and the other made of two new chains. Such a structure has been proposed (Cavalieri and Rosenberg, 1961a,b). This structure, consisting of a *pair of duplexes,* would be thought to arise from another unspecified kind of replication scheme. The essential feature of this proposal is that it requires the biologically produced density hybrid to be of four polynucleotide chains rather than two. These four chains are pictured as being arranged in two conventional duplexes which are united by frequent inter-duplex links. After one round of replication this structure is pictured as one old and one new duplex, whereas the conventional interpretation argues that a single duplex now consists of one old and one new polynucleotide chain. To accept this hypothesis would demand many adjustments in the interpretation of a wide range of experiments, and a full discussion would be quite lengthy, and largely hypothetical.

If the biologically produced hybrid molecules are of this double-duplex structure, then it seems clear that the bonds which unite the duplex strands must have almost the same thermal stability as the hydrogen bonds that unite the polynucleotide chains, for the experiments of Schildkraut, Marmur, and Doty (1961) show that the biologically produced hybrid is converted uniformly with time to heavy and light subunits as the native structure of the molecule (as evidenced by density) is destroyed. Since a 20-minute heating at 90.2°C fails to result in the separation of subunits, while an equivalent heating at 93.8°C does, we must conclude that the hypothetical inter-duplex linkages must be as resistant to heat treatment as the hydrogen bonds, but at the same time no more resistant.

Another special condition must be placed upon these hypothetical linkages between duplex molecules, namely, that of species specificity. DNA samples from different species cannot be renatured together to form molecules of hybrid density, whereas those from the same or a closely related species can (Schildkraut *et al.*, 1961). If the double duplex structure did exist before denaturation, it must have recovered during the renaturation process because transforming DNA has a similar absorbance-temperature profile, and the same rate of thermal inactivation of biological markers as the original DNA, facts which strongly indicate that the basic structure of the molecules is the same before heat denaturation and after renaturation (Marmur and Doty, 1961).

Thus, it is still possible to think of the biologically produced hybrid in terms of a doubled-duplex molecule, provided that one is willing to allow the inter-duplex linkages to have very special properties which are almost the same as the hydrogen bonds which unite the polynucleotide chains. On the other hand, all of the above experiments are easily understood in terms of a biologically produced hybrid molecule made of a single duplex consisting of one new and one old polynucleotide chain.

### C. THE STRUCTURE OF DNA LIBERATED FROM BACTERIA

The great sensitivity of bacteriophage DNA to mechanical breakage by shear means that one should proceed with caution in the interpretation of any molecular weight determined by whatever method as representing a structural or functional subunit of the bacterial genome. The inadvertent fragmentation by stirring or shaking already has proved to be a stumbling block in the extraction of phage DNA; it will be a more acute problem with the extraction of bacterial DNA's which in principle at least could be 50 to 100 times longer than T2 phage DNA molecule. Among the recent investigators who have been studying the extraction of high molecular weight DNA from bacteria one has obtained preparations with sedimentation coefficients of 22 to 29 S corresponding to molecular weights greater than 8 million (Marmur, 1961), while others find 1.2 to 2.4 million, which is thought to be the molecular weight of DNA subunits *in vivo* (Cavalieri *et al.*, 1961).

An interesting kind of DNase has been reported by Bernardi *et al.* (1960, 1961), Bernardi and Sadron, 1961) which reduces the apparent light-scattering molecular weight to about 500,000 and no further. If this proves to be the enzymatic cleavage of unique regions of the DNA molecule, it will prove a useful tool for the examination of DNA organization.

### D. THE GENETIC STRUCTURE OF BACTERIA

The study of bacterial conjugation has led to two major ideas: (*1*) that there does exist a physical structure called the bacterial chromosome which can be transferred from a donor to a recipient cell and which contains all of the genes that have so far been recognized, an (*2*) that there can exist certain strains that contain genetic elements (which contain DNA), which are *not* physically united with the chromosome; these independent elements, which can exist in a number of different forms, are called episomes. Jacob and Wollman (1957, 1958a,b, 1961) have suggested that the bacterial chromosome is originally a circular structure which is broken by the introduction of the fertility factor ($F^+$) to produce the various Hfr strains which are capable of transferring genes (by conjugation) with high frequency to the recipient cell. By examining the order and time sequence of the entry of a number of genes, it was found that each Hfr strain had the same gene order, but a different starting point. Such a situation might arise if these Hfr strains were formed by the fracture of the circular chromosome at various points. This picture of the circular chromosome was greatly strengthened by the discovery of certain Hfr strains which had a *very* high frequency of transfer (Vhf) of the donor chromosome (Taylor and Adelberg, 1960). These strains are able to transfer the *entire* male genome to the female cells in approximately 111 minutes. Using three different Vhf strains, these authors were able to demonstrate more conclusively the continuous linkage of all known chromosomal markers and identify more precisely the separation (in minutes) between markers distributed around the map. In Fig. 11 their map determined by kinetic experiments is shown.

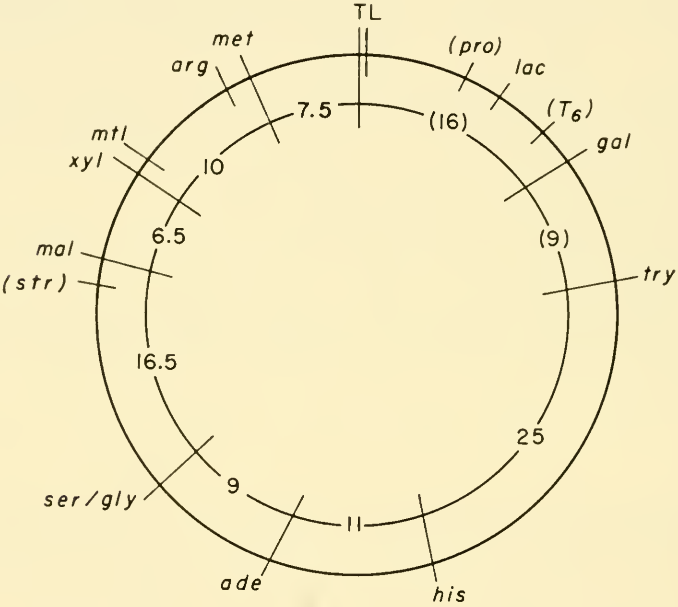


FIG. 11. The kinetic linkage map of *E. coli*. The map represents the circular linkage group of the $F^+$ ancestor of three Vhf strains. Total injection requires about 111 minutes; the various markers enter sequentially and at intervals shown (in minutes). (From Taylor and Adelberg, 1960.)

The genetics of episomes is a subject of some beauty and detail. The reader is referred to the excellent review by Campbell (1962). An important kind of episome is the $F^+$ unit which may be united with any of a number of different segments. This episome may be transferred from a male to a female bacteria by a process called *sexduction* which is analogous to the transfer of the entire chromosome just mentioned (Jacob *et al.*, 1960). This transfer can be accomplished to cells other than coli and in this way *Serratia* cells can be infected with a coli episome (Carey *et al.*, 1960). When *Serratia* cells bearing the *E. coli F-lac* episome are cultured and the DNA is examined in the CsCl density gradient, a small amount of DNA which has a density corresponding to *E. coli* DNA now appears as a small peak close to the peak which is the majority of the *Serratia* DNA. This small amount of DNA which has the density and presumably the composition of *E. coli* DNA is absent in uninfected *Serratia* DNA (Fig. 12A) (Marmur *et al.*, 1961).

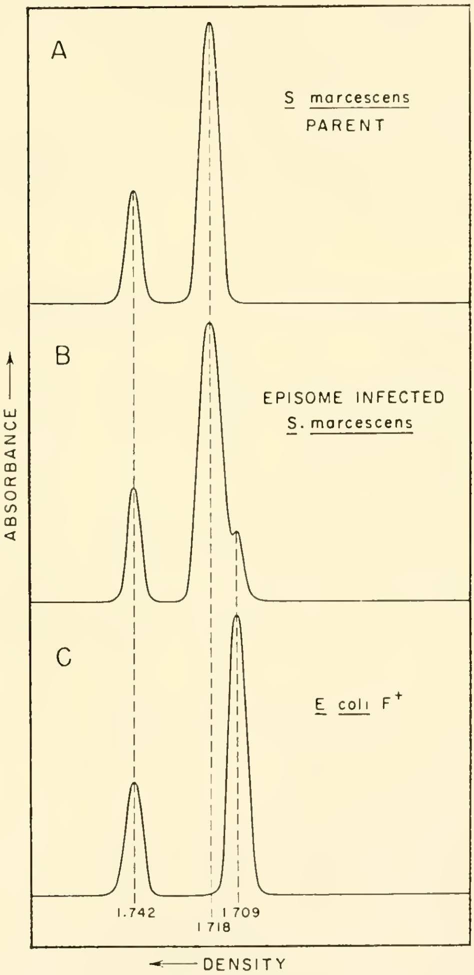


FIG. 12. Band profiles of unfractionated DNA from parental and recombinant strains. Microdensitometer tracings of ultraviolet absorption photographs of samples at equilibrium in a CsCl density gradient at 44,770 rpm. The ordinate represents DNA concentration as a function of the distance from the axis of rotation. The band of buoyant density 1.742 gm/cm$^3$ (N$^{15}$-labeled *Pseudomonas aeruginosa* DNA) is used as a reference standard. (From Marmur *et al.*, 1961.)

## E. CONCLUSIONS

At the present time we can say very little about the physical structure of the DNA in bacteria. Genetically the genome appears circular, but this inference is based on the finding of many different Hfr strains which themselves have a linear genetic structure. The genetic evidence, which has led to the idea of independent genetic elements called episomes, seems to be confirmed by the first density gradient equilibrium experiments.

## ACKNOWLEDGMENTS

The author has profited by many helpful conversations with members of the Biophysics Department, Johns Hopkins University. He is particularly appreciative of the help and criticism of Drs. Irwin Rubenstein and Michael Beer in the preparation of this manuscript.

## REFERENCES

Ames, B. N., Dubin, D. T., and Rosenthal, S. M. (1958). *Science* **127**, 814.

Anderson, E. S., Armstrong, J. A., and Niven, J. S. F. (1959). *9th Symposium Soc. Gen. Microbiol.* p. 224.

Arber, W. (1958). *Arch. sci.* (*Geneva*) **11**, 259.

Aurisicchio, S., Chiozzotto, A., Coppo, A., Donini, P., Frontali, C., and Graziosi, F. (1960). *Nuovo cimento* **18**, Suppl. 2, 197.

Beer, M. (1961). *J. Mol. Biol.* **3**, 263.

Beer, M. (1962). Personal communication.

Beer, M., and Zobel, R. (1961). *J. Mol. Biol.* **3**, 717–726.

Bendet, I., Goldstein, D. A., and Lauffer, M. A. (1960). *Nature* **187**, 781.

Bendet, I., Schachter, E., and Lauffer, M. A. (1962). *J. Mol. Biol.* **5**, 76.

Bernal, J. D., and Fankuchen, I. J. (1941). *J. Gen. Physiol.* **25**, 111.

Bernardi, G., Champagne, M., and Sadron, C. (1960). *Nature* **188**, 228.

Bernardi, G., Champagne, M., and Sadron, C. (1961). *Biochim. et Biophys. Acta* **49**, 1.

Bernardi, G., and Sadron, C. (1961). *Nature* **191**, 809.

Berns, K. I., and Thomas, C. A., Jr. (1961). *J. Mol. Biol.* **3**, 289.

Burgi, E., and Hershey, A. D. (1961). *J. Mol. Biol.* **3**, 458.

Cairns, J. (1962). *J. Mol. Biol.* **3**, 756.

Campbell, A. (1962). *Advances in Genet.* **11**, 101.

Carey, W., Spilman, W., and Baron, L. S. (1960). *Bacteriol. Proc.* p. 73.

Cavalieri, L. F., and Rosenberg, B. H. (1961a). *Biophys. J.* **1**, 317.

Cavalieri, L. F., and Rosenberg, B. H. (1961b). *Biophys. J.* **1**, 323.

Cavalieri, L. F., and Rosenberg, B. H. (1961c). *Biophys. J.* **1**, 337.

Cavalieri, L. F., Deutsch, F., and Rosenberg, B. H. (1961). *Biophys. J.* **1**, 301.

Celano, A., Aurisicchio, S., Coppo, A., Donini, P., and Graziosi, F. (1960). *Nuovo cimento* **18**, Suppl. 2, 190.

Cohen, S., and Barner, H. D. (1954). *Proc. Natl. Acad. Sci. U.S.* **40**, 885.

Cordes, S., Epstein, H. T., and Marmur, J. (1961). *Nature* **191**, 1097.

Davison, P. F. (1959). *Proc. Natl. Acad. Sci U.S.* **45**, 1560.

Davison, P. F. (1962). Personal communication.

Davison, P. F., and Freifelder, D. (1961). Personal communication.

Davison, P. F., and Freifelder, D. (1962). Unpublished experiments.

Davison, P. F., Freifelder, D., Hede, R., and Levinthal, C. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 1123.

Doermann, A. H. (1961). Personal communication.

Doermann, A. H., and Boehner, L. (1962). In press.

Dunn, D. B., and Smith, J. D. (1959). *4th Intern. Congr. Biochem., Vienna, 1958,* p. 72.

Edgar, R. S. (1957). Ph.D. Thesis, University of Rochester, Rochester, New York.

Edgar, R. S. (1961). Personal communication.

Falkow, S., Marmur, J., Carey, W., Spilman, W., and Baron, L. S. (1961). *Genetics* **46**, 703.

Fiers, W., and Sinsheimer, R. L. (1962). *6th Ann. Meeting Biophys. Soc., St. Louis.* Abstr.

Forro, F., Jr., and Wertheimer, S. A. (1960). *Biochim. et Biophys. Acta* **40**, 9.

Fraenkel-Conrat, H. (1962). Personal communication.

Fraenkel-Conrat, H., and Singer, B. (1959). *Biochim. et Biophys. Acta* **33**, 359.

Fraenkel-Conrat, H., and Williams, R. C. (1955). *Proc. Natl. Acad. Sci. U.S.* **41**, 690.

Franklin, R. E. (1955). *Biochim. et Biophys. Acta* **18**, 313.

Fraser, D., and Williams, R. C. (1953). *Proc. Natl. Acad. Sci. U.S.* **39**, 750.

Freese, E. (1958). *Cold Spring Harbor Symposia Quant. Biol.* **23**, 13.

Gellert, M. (1961). *Intern. Biophys. Congr., Stockholm,* p. 288. Abstr.

Gellert, M. (1962). Personal communication.

Graziosi, F. (1961). *Intern. Biophys. Congr., Stockholm,* p. 292. Abstr.

Hamilton, L. D., Barclay, R. K., Wilkins, M. H. F., Brown, G. L., Wilson, H. R., Marvin, D. A., Ephrussi-Taylor, H., and Simmons, N. S. (1959). *J. Biophys. Biochem. Cytol.* **5**, 397.

Hanawalt, P. C., Maaløe, O., Cummings, D. J., and Schaechter, M. (1961). *J. Mol. Biol.* **3**, 156.

Hershey, A. D., and Burgi, E. (1956). *Cold Spring Harbor Symposia Quant. Biol.* **21**, 91.

Hershey, A. D., and Burgi, E. (1960). *J. Mol. Biol.* **2**, 143.

Hershey, A. D., and Chase, M. (1951). *Cold Spring Harbor Symposia Quant. Biol.* **16**, 471.

Hershey, A. D., Burgi, E., and Ingraham, L. (1962). In press.

Jacob, F., Schaeffer, P., and Wollman, E. L. (1960). *10th Symposium Soc. Gen. Microbiol.* p. 67.

Jacob, F., and Wollman, E. (1957). *Compt. rend. acad. sci.* **245**, 1840.

Jacob, F., and Wollman, E. (1958a). *Symposium Soc. Exptl. Biol.* **12**, 75.

Jacob, F., and Wollman, E. (1958b). *Compt. rend. acad. sci.* **247**, 154.

Jacob, F., and Wollman, E. (1961). "Sexuality in Bacteria." Academic Press, New York.

Kahn, P. (1961). Ph.D. Thesis, Yale University, New Haven, Connecticut.

Kaiser, A. D. (1962). *J. Mol. Biol.* **4**, 275.

Kaiser, A. D., and Hogness, D. S. (1960). *J. Mol. Biol.* **2**, 392.

Kellenberger, E. 1960. *10th Symposium Soc. Gen. Microbiol.* pp. 39–66.

Kellenberger, E. (1962). *Advances in Virus Research* **8**, 2.

Kellenberger, E., Sechaud, J., and Ryter, A. (1959). *Virology* **8**, 478.

Kellenberger, G., Zichichi, M. L., and Weigle, J. (1961a). *Proc. Natl. Acad. Sci. U.S.* **47**, 869.

Kellenberger, G., Zichichi, M. L., and Weigle, J. (1961b). *J. Mol. Biol.* 3, 399.

Kleinschmidt, A., Gehatia, M., and Zahn, R. K. (1960). *Kolloid-Z.* 169, 156.

Klug, A., and Casper, D. L. D. (1960). *Advances in Virus Research* 7, 225.

Koch, G., and Hershey, A. D. (1959). *J. Mol. Biol.* 1, 260.

Kozinski, A. W. (1961). *Virology* 13, 124.

Kozinski, A. W. (1962). Personal communication.

Langridge, R., Seeds, W. E., Wilson, H. R., Hooper, C. W., Wilkins, M. H. F., and Hamilton, L. D. (1957). *J. Biochem. Biophys. Cytol.* 3, 767.

Langridge, R., Seeds, W. E., Wilson, H. R., Hooper, C. W., Wilkins, M. H. F., and Hamilton, L. D. (1960a). *J. Mol. Biol.* 2, 19.

Langridge, R., Seeds, W. E., Wilson, H. R., Hooper, C. W., Wilkins, M. H. F., and Hamilton, L. D. (1960b). *J. Mol. Biol.* 2, 38.

Lehman, I. R. (1960). *J. Biol. Chem.* 235, 1479.

Lerman, L. J. (1961). *J. Mol. Biol.* 3, 18.

Levine, L., Barlow, J. L., and Van Vunakis, H. (1958). *Virology* 6, 702.

Levine, L., Murakami, W. T., Van Vunakis, H., and Grossman, L. (1960). *Proc. Natl. Acad. Sci. U.S.* 46, 1038.

Levinthal, C. (1954). *Genetics* 39, 169.

Levinthal, C. (1956). *Proc. Natl. Acad. Sci. U.S.* 42, 394.

Levinthal, C., and Davison, P. (1961). *J. Mol. Biol.* 3, 674.

Levinthal, C., and Thomas, C. A., Jr. (1957). *Biochim. et Biophys. Acta* 23, 453.

Luzzati, V., Nicolaieff, A., and Masson, F. (1961). *J. Mol. Biol.* 3, 185.

Maaløe, O. (1962). Personal communication.

Maaløe, O., and Hanawalt, P. C. (1961). *J. Mol. Biol.* 3, 144.

Maestre, M. F., and Kilksom, R. (1962). *Nature,* in press.

Mahler, H. R., and Fraser, D. (1961). *Nature* 189, 948.

Mandell, J. D., and Hershey, A. D. (1960). *Anal. Biochem.* 1, 66.

Marmur, J. (1961). *J. Mol. Biol.* 3, 208.

Marmur, J., and Doty, P. (1961). *J. Mol. Biol.* 3, 585.

Marmur, J., Rownd, R., Falkow, S., Baron, L. S., Schildkraut, C., and Doty, P. (1961). *Proc. Natl. Acad. Sci. U.S.* 47, 972.

Matheson, A. T., and Thomas, C. A., Jr. (1960). *Virology* 11, 289.

Marvin, D. A., Spencer, M., Wilkins, M. H. F., and Hamilton, L. D. (1961). *J. Mol. Biol.* 3, 547.

Mayor, H. D., and Hill, N. O. (1961). *Virology* 14, 264.

Meselson, M. (1960). *In* "The Cell Nucleus" (J. S. Mitchell, ed.), p. 240. Academic Press, New York.

Meselson, M., and Stahl, F. W. (1958a). *Proc. Natl. Acad. Sci. U.S.* 44, 671.

Meselson, M., and Stahl, F. W. (1958b). *Cold Spring Harbor Symposia Quant. Biol.* 23, 1.

Meselson, M., and Weigle, J. (1961). *Proc. Natl. Acad. Sci. U.S.* 47, 857.

Minagawa, T. (1961). *Virology* 13, 515.

Nomura, M., and Benzer, S. (1961). *J. Mol. Biol.* 3, 684.

North, A. C. T., and Rich, A. (1961). *Nature* 191, 1242.

Putnam, F. W. (1953). *Advances in Protein Chem.* 8, 175.

Rolfe, R. (1962). *J. Mol. Biol.* 4, 22.

Roller, A. (1961). Ph.D. Thesis, California Institute of Technology, Pasadena.

Rubenstein, I. (1960). Ph.D. Thesis, University of California, Los Angeles.

Rubenstein, I., Thomas, C. A., Jr., and Hershey, A. D. (1961). *Proc. Natl. Acad. Sci. U.S.* 8, 1113.

Schildkraut, C., Marmur, J., and Doty, P. (1961). *J. Mol. Biol.* 3, 595.

Sinsheimer, R. L. (1959a). *J. Mol. Biol.* **1**, 37.

Sinsheimer, R. L. (1959b). *J. Mol. Biol.* **1**, 43.

Sinsheimer, R. L. (1960). *In* "The Nucleic Acids" (E. Chargaff and J. N. Davidson, eds.), Vol. III, p. 187. Academic Press, New York.

Sinsheimer, R. L. (1961). *5th Ann. Meeting Biophys. Soc., St. Louis.* Abstr. TA 9.

Stahl, F. W. (1962). Personal communication.

Stent, G. S., and Fuerst, C. J. (1960). *Advances in Biol. and Med. Phys.* **7**, 2.

Stent, G. S., Sato, G. H., and Jerne, N. K. (1959). *J. Mol. Biol.* **1**, 134.

Streisinger, G., and Bruce, V. (1960). *Genetics* **45**, 1289.

Streisinger, G., Edgar, R. S., and Harrar, G. (1961). Personal communication.

Studier, F. W., and T'so, P. (1961). Personal communication.

Taylor, A. L., and Adelberg, E. A. (1960). *Genetics* **45**, 1233.

Taylor, J. H., Woods, P. S., and Hughes, W. L. (1957). *Proc. Natl. Acad. Sci. U.S.* **43**, 122.

Tessman, I. (1959). *Virology* **7**, 263.

Tessman, I., Tessman, E. S., and Stent, G. S. (1957). *Virology* **3**, 209.

Thomas, C. A., Jr. (1959). *J. Gen. Physiol.* **42**, 502.

Thomas, C. A., Jr., and Berns, K. I. (1961). *J. Mol. Biol.* **3**, 277.

Thomas, C. A., Jr., and Berns, K. I. (1962). *J. Mol. Biol.*, in press.

Thomas, C. A., Jr., and Pinkerton, T. C. (1962). *Gen. Mol. Biol.* **5**, 356.

Trautner, T. A. (1958). *Z. Verbungslehre* **89**, 264.

Van Tubergen, R. P., and Setlow, R. B. (1961). *Biophys. J.* **1**, 589.

Weigle, J., Meselson, M., and Paigen, K. (1959). *J. Mol. Biol.* **1**, 379.

Wilkins, M. H. F., Gosling, R. G., and Seeds, W. E. (1951). *Nature* **167**, 759.

## *Chapter* IV

# Cellular Control of DNA Biosynthesis

KARL G. LARK

*Department of Microbiology,*
*St. Louis University School of Medicine,*
*St. Louis, Missouri*

I. Introduction . . . . . . . . . . . . . . . . . . . 153
  A. The DNA Cycle . . . . . . . . . . . . . . . . . 155
  B. The Level of Biochemical Control . . . . . . . . . . 160
II. Role of DNA Precursors in Controlling DNA Synthesis . . . . 162
  A. The Direct Measurement of Deoxynucleotides or Precursors Thereof and Their Relation to the DNA Cycle . . . . . . 163
  B. The Measurement of Enzyme Activity Associated with the Production of DNA Precursor . . . . . . . . . . 170
III. Role of Protein or RNA Synthesis in Controlling DNA Synthesis . . 173
IV. The Physical State of DNA as a Factor Controlling Its Replication . . 180
  A. Is It Necessary for DNA to Exist in a Special Form in Order for It to Be Duplicated? . . . . . . . . . . . . . . . 183
  B. Is DNA Synthesis Compatible with Simultaneous RNA Synthesis? . 186
  C. Is DNA Structurally Stabilized within the Cell and Is This a Necessary Part of DNA Replication? . . . . . . . . . 188
  D. Is the Replication of DNA Physically Ordered, with Respect to the Macromolecule Itself, with a Beginning, a Middle, and an End? . 188
V. Other Factors Affecting DNA Biosynthesis and the DNA Cycle . . 190
  A. The Localization of DNA Synthesis . . . . . . . . . 190
  B. Variation in the DNA Cycle . . . . . . . . . . . . 191
  C. Effects of Temperature and Radiation on the DNA Cycle . . . 192
  D. The Effect of the Nutrient Environment . . . . . . . . 194
  E. The Dissociation of the Cell Division Cycle from DNA Synthesis . 197
VI. Summary, Conclusion, and Opinions . . . . . . . . . . 197
  Acknowledgments . . . . . . . . . . . . . . . . . 200
  References . . . . . . . . . . . . . . . . . . . 200

## I. INTRODUCTION

The biosynthesis of DNA is a distinctive characteristic of cell proliferation rather than of the broad span of phenomena commonly associated with actively functioning cells. Cells which are rapidly metabolizing may often exhibit little or no DNA synthesis, whereas division requires a previous replication of genetic material if competent daughter cells are to be produced. The mechanisms controlling the biosynthesis of DNA are of interest in the study of problems of differentiation as well as those studies concerned with the maintenance of genetic continuity, since absence of DNA synthesis may prevent, and thus control, further cell division without impairing cell function.

Individual cells possess the ability to control DNA synthesis. During a normal cycle of cellular division the content of DNA, as measured chemically and genetically, is *exactly* doubled. The existence of this situation has been inferred for some time from the knowledge that during multiplication the DNA content of individual cells of a given species or clone remains remarkably constant and that the chromosome number of the vast majority of individuals arising in a clone also does not increase or decrease. Instances in which a dividing cell divests itself of genetic material as a result of asynchrony between cell division and duplication of DNA are the exception rather than the rule; cells in which DNA synthesis and cell division have ceased usually resume DNA synthesis before division recommences.

The extent to which intracellular DNA synthesis is controlled has become apparent in recent years, through use of cellular autoradiography, quantitative microspectrophotometry, and synchronous culture techniques. As a result of these studies it has become evident that most animal, plant, or protozoan cells, as well as some bacteria, have a defined period during their life cycle during which DNA is synthesized.

The living cell therefore exhibits two types of control over the biosynthesis of DNA: one, quantitative; the other, temporal. The amount synthesized is always equal to that initially present and the time of synthesis is restricted to a defined interval during the division cycle.

Although the evidence for this control of DNA synthesis has been available for some time, few systematic attempts have been made to understand the mechanism of control which governs the "DNA cycle." This has been due, in part, to a natural desire for further evidence to substantiate the generality of such a cycle and, in part, to a lack of knowledge of the biochemical mechanisms underlying the synthesis of DNA. Until recently, microscopic observation was found to be of little use in studying the interphase nucleus, in which DNA is synthesized.

In recent years the enzymatic basis for DNA synthesis has become known, and work has concentrated upon the study of the *in vivo* replication mechanism of this genetic material. In addition, physicochemical information on the structure of the molecule, spurred on by the development of new physical techniques and genetic concepts, has led to a detailed knowledge of its structure.

Finally, a more detailed knowledge of the mechanisms whereby DNA controls the synthesis of cellular material has led to an ever-increasing interest in the role of DNA in regulating the cell's physiological behavior.

A large amount of information is now available relating to the control of DNA biosynthesis, but little of it has been obtained with this end in view. In what follows, material from different sources, investigated originally in a variety of contexts, is discussed. This will lead to the inevitable criticism that results are not strictly comparable and that, therefore, general conclusions may not be drawn. However, what follows presents a definition of the problem, the current status of its investigation, and the direction in which it appears that further investigation can profitably be directed.

It is intended to represent one individual's analysis of the situation rather than a comprehensive review of the literature.

### A. THE DNA CYCLE

Recent studies have led to the general conclusion that during the division cycle of higher cells there exists a period following mitosis during which DNA synthesis does not take place ($G_1$), a subsequent period of DNA synthesis during which the DNA content of the interphase nucleus is doubled ($S$), a second period during which DNA is not synthesized ($G_2$), and, finally, a period of mitosis ($M$) (Howard and Pelc, 1953) (Fig. 1). Although the duration of these periods may vary from system to system, they are often remarkably constant from cell to cell within the same system. [In bacteria, it has been possible to demonstrate that DNA is synthesized throughout most of the cell cycle (Schaechter *et al.*, 1959; Young and Fitz-James, 1959; McFall and Stent, 1959; Abbo and Pardee, 1960). However, recent evidence has indicated that even in these systems a DNA duplication cycle may exist

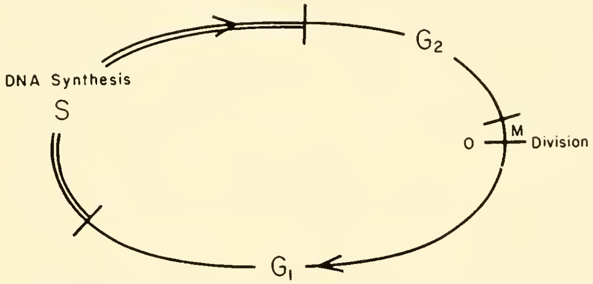


FIG. 1. A representation of the DNA synthetic cycle in relation to the cell division cycle. Division, or the end of mitosis, marks the beginning of the cycle. A gap, $G_1$, occurs before the onset of DNA synthesis, $S$. DNA synthesis is followed by another period, $G_2$, prior to the onset of mitosis. (Adapted from Stanners and Till, 1960.)

(Maaløe and Hanawalt, 1961; Hanawalt *et al.*, 1961; Goldstein *et al.*, 1959). It seems possible, therefore, that such systems may represent situations in which $G_1$ and $G_2$ may be of extremely short duration relative to $S$. This is made more probable by the finding that $G_1$ or $G_2$ may be extended in such systems by simple chemical or physical treatments which also synchronize the cell population (Lark and Maaløe, 1956; Barner and Cohen, 1956; Maruyama, 1956; Scott and Chu, 1958; Maruyama and Lark, 1959).]

The existence of a cycle such as that in Fig. 1 has been demonstrated with autoradiography (Howard and Pelc, 1953; Lajtha *et al.*, 1954; Hornsey and Howard, 1956; Taylor *et al.*, 1957; Firket and Verly, 1958; Mendelsohn *et al.*, 1960; Painter and Drew, 1959; Stanners and Till, 1960; Prescott, 1960; Edwards *et al.*, 1960; Sisken and Kinosita, 1961) by labeling cells at a given time with a radioactive precursor of DNA (usually tritiated thymidine) and then measuring the period of time required for the first labeled mitotic figures to appear subsequently as well as the proportion of cells so labeled or the rate at which radioactivity, incorporated during interphase, appears in metaphase mitotic figures. These measurements yield the time elapsing between the end of

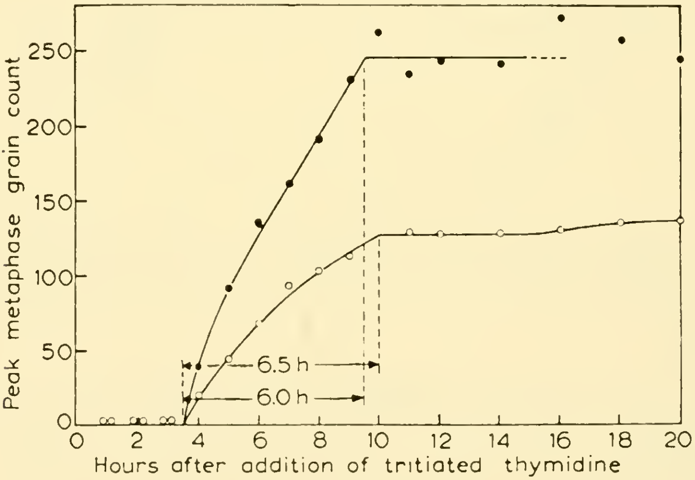


FIG. 2. The peak metaphase grain count values from autoradiographs of cells sampled at various times after addition of thymidine-$H^3$ to L-strain mouse cells growing in tissue culture. The upper curve represents a film exposure of 15 days; the lower, 8. Peak counts for each sample were obtained using the peak grain count values from a frequency distribution of the grain count for each sample. (Number of metaphases as a function of their individual grain count.) $G_2$ is the period before an increase in grain count is observed. The grain count rises throughout $S$, after which it is maintained at a constant value. (From Stanners and Till, 1960.)

DNA synthesis and the beginning of mitosis as well as the duration of the period of DNA synthesis. (For a detailed review of the design and interpretation of such experiments see Taylor, 1960a.)

Figure 2 is an example of such an experiment in which exponential multiplying L cells (tissue culture strain originally derived from the mouse) were labeled with thymidine-$H^3$ and the peak grain count at metaphase measured at intervals thereafter. As may be seen, no counts were recorded for 3–4 hours after the onset of labeling; this is the period ($G_2$) between the cessation of DNA synthesis and metaphase. The number of grains/metaphase nucleus then rises for 6–7 hours after which no increase in labeling is observed; this, then, is the period of DNA synthesis, or $S$, during which incorporation of thymidine can occur. The number of grains will be low for those nuclei which were leaving $S$ when thymidine-$H^3$ was introduced. The grain number per nucleus will rise until a maximum value is reached which represents nuclei which have passed through their entire $S$ period in the presence of the radioactive isotope. The duration of $S$ is thus given by the length of time during which the grain count/nucleus rises. Since the generation time of these cells was 20 hours and mitosis occupies a period of 1/3 hour, $G_1$ may be estimated as between 9 and 11 hours. Note the sharply defined appearance of labeled mitoses and of the saturation of the grain count, indicating the similar behavior of different individuals of the random population.

The cycle of DNA synthesis may also be reconstructed from microspectrophotometric measurements of the DNA content of individual cells sampled at random (Swift, 1950; Walker and Yates, 1952; Alfert, 1955; Bloch and Godman, 1955; Vendrely and Vendrely, 1956; Richards *et al.*, 1956; McDonald, 1958; Walker and Mitchison, 1958; Kimball and Borka, 1959; Smith *et al.*, 1959; Painter *et al.*, 1960; McLeish and Sunderland, 1961; Utsume, 1961). Such reconstructions have been substantiated by determination of the DNA content of individual cells at known intervals following mitosis and by measurements of synchronous cultures.

Figure 3 represents an example of such an experiment using embryonic mouse heart cells grown in tissue culture. Two observations are noteworthy: (*a*) the similarity between the DNA cycle obtained by this method and that illustrated in Fig. 2, and (*b*) the precise control of the quantity of DNA within the cells, the number of cells containing amounts of DNA less than 22 or greater than 44 units being negligible.

Finally, experiments on synchronous cell populations have measured the total amount of DNA in the population and assumed that this would reflect the course of synthesis by a single cell (Ogur *et al.*, 1953; Hotch-

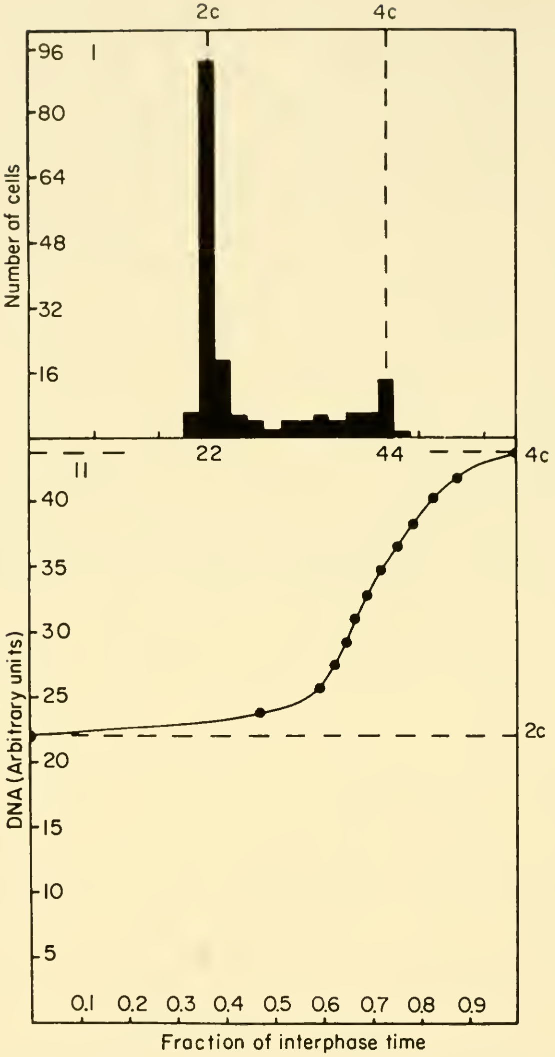


Fig. 3. The DNA content of cells from mouse heart, growing in tissue culture. DNA content was measured microspectrophotometrically on Feulgen-stained cells. The number of cells on the ordinate contain the amount of DNA shown on the abscissa. The lower curve was constructed from this distribution assuming an exponential rate of cell growth and that a constant amount (previously determined) of the division cycle was devoted to interphase. (From Richards *et al.*, 1956.)

kiss, 1954; Young and Fitz-James, 1959; Iwamura, 1955; Barner and Cohen, 1956; Lark and Maaløe, 1956; Maruyama, 1956; Scott and Chu, 1958; Maruyama and Lark, 1959; Foster and Stern, 1959; Newton and

Wildy, 1959; Nygaard *et al.*, 1960; Williamson and Scopes, 1960; Abbo and Pardee, 1960; Burns, 1961; Terasima and Tolmach, 1962). Figure 4 (open circles) shows the course of DNA synthesis as well as of mitosis in dividing microspores of the anther of *Trillium* (Erikson, 1948; Stern, 1960b; Hotta and Stern, 1961b). Chemical analyses were performed on samples of isolated microspores. It should be noted that the amount of DNA synthesized is exactly equal to that initially present.

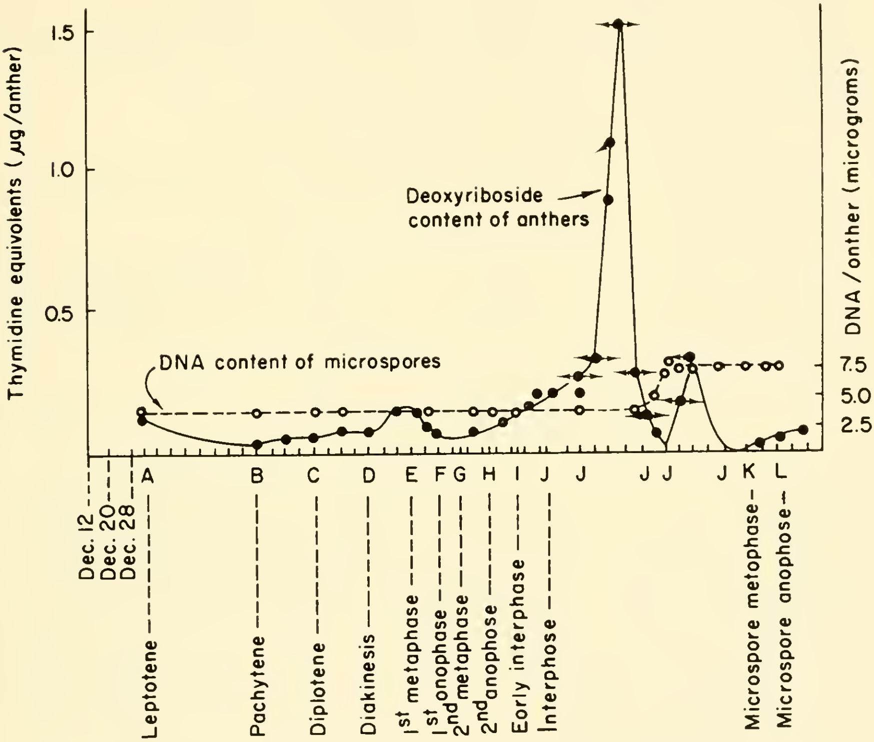


FIG. 4. DNA synthesis in the microspores and deoxyriboside production in the anther of *Trillium*. Synchronous development of the microspores occurs and as a result DNA doubles during a defined period of interphase. A double peak of deoxyriboside production by the anther is observed at this time. The possibility exists that the second peak may represent an overshoot in microspore DNA precursor production (see text, p. 163). (From Hotta and Stern, 1961b.)

All of these experiments serve to demonstrate the temporal control outlined in Fig. 1. A similar situation is sometimes seen within a single cell containing more than one type of nucleus or in a single, very large nucleus. Thus, in the ciliate *Euplotes* (Gall, 1959) synthesis of DNA in the large U-shaped macronucleus (which is a functional nucleus in contrast to the micronucleus which maintains the genetic continuity of the cell) starts *simultaneously* at the ends of the two lobes and proceeds toward the center. Similarly, in paramecia, synthesis of DNA is initiated simultaneously in the macro- and micronuclei, although synthesis is completed in the latter within a much shorter time than the former (Woodard *et al.*, 1961). This suggests that the initiation and rate of DNA synthesis in both the macro- and micronucleus may be controlled by a similar mechanism. (However, see Section V,A below.)

The above examples, though only a few of the large number to be found in the literature, illustrate the type of control in which we are interested.

### B. THE LEVEL OF BIOCHEMICAL CONTROL

From our present biochemical knowledge, what possibilities exist which could serve as a basis for a control mechanism?

Figure 5 is a diagrammatic representation of the sequence of biochemical steps in the synthesis of DNA. The experimental basis for this sequence has been discussed extensively in the first chapter of this book. Mechanisms for controlling the biosynthesis of DNA may, in theory, be encountered at several levels.

The requirement for deoxynucleotide triphosphates as the substrate for polymerization makes the production of these precursors an important point of control in the synthesis of DNA (Bessman *et al.*, 1958). (Control, at a level less specific than this, would simultaneously affect other synthetic processes such as RNA and protein synthesis. For this reason only more specific control mechanisms will be considered.) The synthesis of these molecules, in turn, may be controlled at:

(*a*) the conversion of ribonucleotides to deoxyribonucleotides (Reichard *et al.*, 1961) and the formation of thymidylic acid (Friedkin, 1959), and/or

(*b*) the phosphorylation of deoxyribonucleotides to the triphosphate form (Lehman *et al.*, 1958).

Thus, the rates at which the triphosphates are formed are governed by several enzyme reactions subject to control through feedback inhibition or repression (Pardee, 1961). Alternatively, competitive factors may be operative; for example, a precursor substrate such as a ribonucleotide may have been depleted by the synthesis of compounds other than DNA.

The actual polymerization of DNA presents a variety of opportunities for the operation of a control mechanism. This reaction requires the participation of the "old" or pre-existent DNA which acts as both a

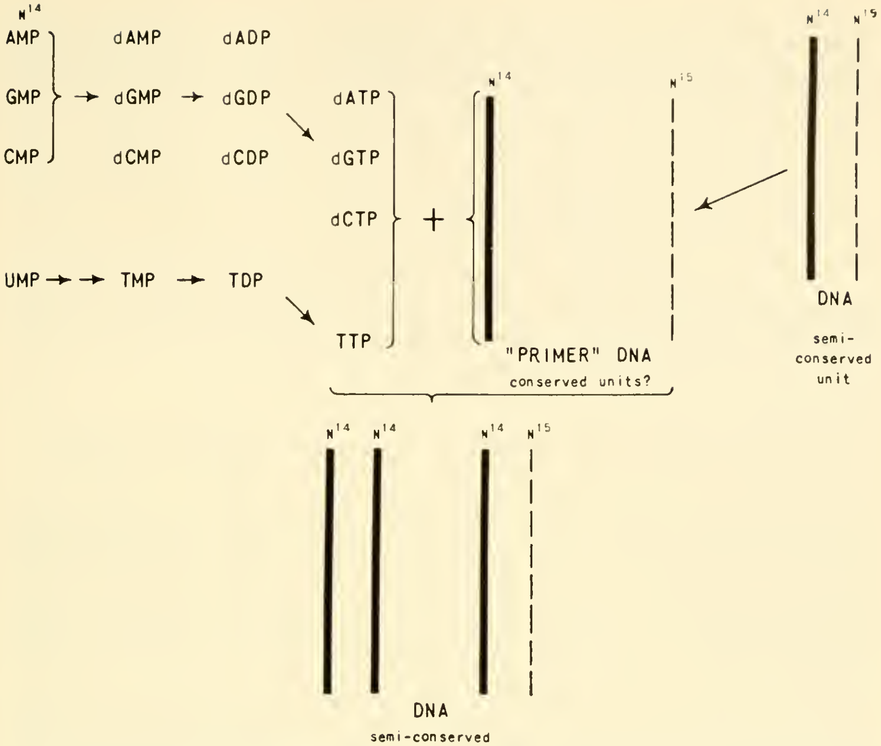

$N^{14}$
AMP
GMP
CMP
dAMP
dGMP
dCMP
dADP
dGDP
dCDP
UMP
TMP
TDP
dATP
dGTP
dCTP
TTP
+
$N^{14}$
$N^{15}$
"PRIMER" DNA
conserved units?
$N^{14}$
$N^{15}$
DNA
semi-conserved unit
$N^{14}$
$N^{14}$
$N^{14}$
$N^{15}$
DNA
semi-conserved

FIG. 5. Schematic representation of the biochemical sequence of events involved in the biosynthesis of DNA. The ribonucleoside monophosphates—adenylic acid (AMP), cytydilic acid (CMP), guanylic acid (GMP), or uridylic acid (UMP)—are converted to deoxyribotides dAMP, dGMP, dCMP, or TMP (thymidylic acid) which are, in turn, converted to the di- and triphosphate derivatives (dADP, dGDP, dCDP, TDP, and dATP, dGTP, dCTP, and dTTP). These are polymerized under the direction of "primer" DNA to form new DNA. The primer DNA is derived from pre-existing DNA and may represent the conserved unit of this semi-conserved molecule. Thus, if the pre-existent DNA carries a semi-conserved density label such as $N^{15}$, the product DNA will carry a density label distribution (as shown), demonstrating the semi-conservative nature of DNA replication.

primer and a template in the formation of new material (Kornberg, 1959). In doing this, the old and new polynucleotide strands are not confused; there is one for one pairing of strands and these remain as separate entities which may be distinguished by a variety of physical techniques (Levinthal, 1956; Taylor *et al.*, 1957; Meselson and Stahl, 1958; Mendelsohn, 1960; Sueoka, 1960; Forro, 1960; Simon, 1961). This "semi-conservative" method of replication appears to extend to aggregates of DNA molecules which may be as large as chromosomes (the evidence for which has been discussed in Chapter II).

By entering a primer state DNA may initiate its own replication which may again be terminated when primer is no longer available. During DNA replication the conversion to primer may be a complicated process which may involve structural changes in the molecule itself and in its association with other molecules such as protein and RNA.

At this level, then, control of DNA biosynthesis may involve the synthesis of RNA or protein or both.

Finally, it should be realized that restrictions on the proximity of the various components of the polymerizing reaction may play an important role. Thus, if the various enzymes involved in DNA synthesis (kinase and polymerase) are formed in the cytoplasm, they must enter the nucleus in order to utilize the DNA primer.

## II. Role of DNA Precursors in Controlling DNA Synthesis

In order to synthesize DNA the cell must produce the necessary precursor deoxyribotides. *In vitro* enzymatic studies have shown that these must be present in the triphosphate form and that all of the nucleotides found in the DNA which is being synthesized must be present simultaneously. There is every reason to believe that the results of these studies are valid *in vivo* as well as *in vitro*.

For example, bacteria, upon infection with bacteriophage, commence the rapid production of large quantities of viral DNA which continues until the cell lyses (Cohen, 1947; Hershey *et al.*, 1953). It has been shown that a primary step in the conversion of the cell into such a "DNA factory" is the production of many new enzymes needed for the synthesis of new types of phage-specific deoxynucleotides (Flaks and Cohen, 1957; Kornberg *et al.*, 1959), as well as the production of increased quantities of pre-existing enzymes also associated with DNA synthesis. All of these are necessary for producing the deoxynucleotide triphosphates required for DNA synthesis. As would be expected, such phage-infected cells contain large quantities of acid-soluble purine and pyrimidine deoxynucleotides, whereas uninfected cells contain small amounts of only the pyrimidine deoxynucleotides (O'Donnell *et al.*, 1958). Moreover, to further enhance the efficiency of the system the enzymatic balance of the cell is shifted to emphasize the synthesis of DNA rather than RNA (Cohen, 1947). Although this system may be regarded as one in which DNA synthesis continues "out of control" it exemplifies a biological system in which the cell follows the predictions of the enzyme chemists, with respect to precursor production and DNA biosynthesis.

Granted that the production of DNA precursors is a necessary step in DNA synthesis, to what extent does such precursor production control DNA biosynthesis during the division cycle of the normal cell?

To date, experiments to answer this question have utilized synchronized cell systems of various types. Two types of experimentation may be considered:

### A. THE DIRECT MEASUREMENT OF DEOXYNUCLEOTIDES OR PRECURSORS THEREOF AND THEIR RELATION TO THE DNA CYCLE

The most striking example of a control mechanism which may rest on precursor production is that of DNA synthesis in the microspores of the anthers of *Lilium longiforum* and *Trillium erectum*. The lily system has proved useful because of the fact that the length of the flower bud enclosing the anthers is correlated with the synchronized activities of the microspore (Erikson, 1948). In *Trillium,* microspore development is also synchronized with respect to time. By sampling anthers from a given bud at different times, or by selecting buds of the proper length, it is possible to examine large samples of material at known stages of the division cycle and to study the synthesis of DNA by the microspores (Taylor and McMaster, 1954; Foster and Stern, 1958). When this is done, it can be shown that the deoxyribosides of guanine, adenine, cytosine, and thymine appear in the fluid surrounding the microspores shortly before DNA synthesis and mitosis (Stern, 1960a). These disappear abruptly as DNA is made (Fig. 4). It has been shown that such material originates from the destruction of a polydeoxyribotide in extra sporageneous tissue (Stern, 1961) and that the deoxynucleosides thus formed are rapidly converted in the microspores or the microspore walls into triphosphates, the "half-life" of which is very short (Stern, 1960b; Hotta and Stern, 1961a). Recent evidence has demonstrated that the ability of microspores to phosphorylate thymidine increases immediately following the appearance of deoxyribosides in the extracellular fluid. It therefore appears that these phosphorylating enzymes (necessary to produce deoxynucleotide triphosphates) may be induced by deoxyribosides (Hotta and Stern, 1961b). Although it has not been possible to demonstrate directly the conversion of the extracellular deoxyriboside material into DNA, the data at present are consistent with a mechanism by which the presence of deoxyribosides initiate the synthesis of DNA, presumably by making a pool of precursor material available.

One difficulty, however, is that two cycles of deoxyriboside production and disappearance occur for each step of DNA synthesis, a pattern which is most striking in experiments with *Lilium*. This may arise from a slight shift in the phase of synchrony among anthers from a single bud (Stern, 1962) or may indicate a role of deoxyribosides other than as a precursor for DNA (Stern, 1960a). If the former is true, a similar difference should appear in the synthesis of DNA which should then occur as a two-step process. This has not been observed. Moreover, the amount of DNA synthesized is less than the deoxyriboside pool or the amount of DNA broken down to create the pool (this loss of DNA may indicate the total size of the deoxyriboside pool formed) (Stern, 1960a).

These difficulties appear to be explained by the recent work of Takats (1962) who has demonstrated that the deoxyriboside material which is derived from the breakdown of tapetal DNA is *not* used in the synthesis of microspore DNA. He labeled the tapetum with thymidine-$H^3$ at a time when no DNA was being synthesized in the microspores and subsequently traced the breakdown and reutilization of this radioactive material by means of autoradiography. It was found that the degraded material was taken up in the microspore walls but not into microspore DNA. He suggests that this material could serve to induce the formation of a second deoxyriboside or -tide pool (this might explain the double peak of acid-soluble material prior to DNA synthesis) which is then used to supply the necessary precursors for microspore DNA synthesis. This idea is similar to that of Hotta and Stern (1961), who have suggested that this pool of material is responsible for the induction of nucleotide phosphorylating enzymes in the anther.

It would appear, then, that control of DNA synthesis in the lily anther may be the result of a precursor pool *induced* by the products from the degradation of tapetal DNA. The induction of a second pool of deoxyriboside material is further suggested by the finding, by Stern and Hotta (1962), of two types of DNA in cells of the anther wall, one type having the base ratios and properties of the microspore DNA. Breakdown of this material could easily provide a pool of precursors for the synthesis of microspore DNA.

A situation similar to that found in the division of lily microspores may exist during the development of fertilized eggs and embryos of amphibia and Echinodermata. Previous evidence (Hoff-Jørgensen and Zeuthen, 1952; Gregg and Løtrup, 1955; Finamore and Volkin, 1958) had indicated a high extra nuclear (cytoplasmic) content of DNA-like substances in developing amphibian embryos. It was presumed that this material serves as a source of nuclear DNA. It has been shown (Grant, 1958; Kuriki and Okazaki, 1959) that a large portion of this "excess DNA" is composed of deoxynucleotides which are present in amounts far in excess of the nuclear DNA. The presence of all of the pyrimidine and purine deoxynucleotide triphosphates was noted. Thus this material could serve as either a reservoir of precursors for DNA synthesis or as an inducer for the synthesis of nuclear deoxyribotides. No correlative studies were made between the utilization of these compounds and nuclear DNA synthesis.

It has also been shown that sea urchin embryos contain large amounts of acid-soluble deoxyribosidic compounds (Sugino *et al.*, 1960; Sugino, 1960) amounting to more than 5 or 6 times the intracellular DNA content. In these studies it was shown that the DNA content of the cell doubled during the cell division cycle. However, no studies of the time-course of DNA synthesis as related to the disappearance of acid-soluble deoxyribosidic compounds were made. The acid-soluble material appears to be derived from cytidine, which would indicate that it cannot alone serve as an adequate reservoir of DNA precursor material. Nevertheless, it is possible that further, more detailed studies on this system could reveal a relationship similar to that found in amphibian embryos. That this may be the case is indicated by the immunity of echinoderm embryo development to agents inhibiting purine or pyrimidine deoxynucleoside biosynthesis (Karnofsky and Bevelander, 1958; Karnofsky and Basch, 1960). This immunity persists until the blastula stage and indicates that until this time the developing embryo is not dependent on *de novo* purine or pyrimidine biosynthesis. It is possible that further studies in these systems of the time relationship between DNA synthesis and deoxyribotide utilization may substantiate a product-precursor relationship between the two.

In these systems, then, the possibility exists that changes in the quantity of precursor material may control the synthesis of DNA.

Other situations in which deoxyribosides, or deoxyribotides, accumulate quantitatively prior to DNA synthesis are not known. Material from a number of different sources have been analyzed for their content of deoxyribosides or deoxyriboside derivatives (Schneider, 1955, 1957; Schneider and Brownell, 1957; Rotherham and Schneider, 1958, 1960; Potter and Schlesinger, 1955; Potter *et al.*, 1957; Potter and Buettner-Janusch, 1958; LePage, 1957; Ord and Stocken, 1958; Parizek *et al.*, 1958; Davidson and Bishop, 1957; Okazaki and Okazaki, 1958; Okazaki *et al.*, 1959; Okazaki, 1959; O'Donnell *et al.*, 1958; Okazaki *et al.*, 1960; Lark, 1961). With one exception (Kanazir, 1954), the amount of material was only a small fraction (from 0.2–5%, depending upon the system) of the DNA content of the cell. Moreover, with the exception of two observations only pyrimidine deoxynucleosides or deoxynucleotides have been reported. In one of these exceptions where deoxyadenosine triphosphate was observed in a rat carcinoma (LePage, 1957), this amounted to less than 10% of the quantity of pyrimidine deoxynucleosides or deoxynucleotides (Schneider, 1957).

Although a large quantity of precursors might not be found in a population of randomly dividing cells, cells undergoing synchronous division and DNA synthesis might well be expected to show such an accumulation. This is apparently not so. In regenerating liver[1] an increase in deoxynucleoside and deoxynucleotide content is observed but the amounts of these are still extremely small and the quantity present at 24 hours (presumably the onset of DNA synthesis) are not strikingly greater (at most 50%) than the amounts found in hepatomas (Schneider and Brownell, 1957; Potter *et al.*, 1957; Rotherham and Schneider, 1958) or regenerating liver after 48 hours (following DNA synthesis). Thus the increase observed appears consistent with the overall increase in DNA synthetic activity and cell division in this tissue.

Finally, it has not been demonstrated that all of such material serves as precursors for DNA synthesis (Hecht and Potter, 1956; Potter and Buettner-Janusch, 1958). Especially significant is the fact that no purine deoxyriboside material is observed.

On the other hand, in regenerating liver and in hepatomas the proportion of deoxynucleotides is increased relative to nucleosides and blockage of DNA synthesis following irradiation will result in accumulation of deoxypyrimidines in the tissue and excretion of such material in the urine (Ord and Stocken, 1958; Parizek *et al.*, 1958; Davidson and Bishop, 1957; Main *et al.*, 1957). It seems probable, therefore, that these substances *are* acting as precursors (see discussion of Fig. 7 below).

Bacteria have also been found to contain a small pool of deoxypyrimidine compounds (Okazaki and Okazaki, 1958; Okazaki *et al.*, 1959; Okazaki, 1959; O'Donnell *et al.*, 1958; Okazaki *et al.*, 1960; Lark, 1961). To date, almost all of these have been shown to be thymidine derivatives and evidence is accumulating which indicates that they may act as intermediates in the transport or activation of certain sugars (Strominger and Scott, 1959; Okazaki, 1959; Okazaki, 1960; Baddiley and Blumson, 1960). With one exception, no deoxypurine or derivatives thereof have been observed. [In the case of *E. coli* B, Kanazir (1954) has reported the existence of large amounts of deoxyguanosine within the acid-soluble pool. Other workers, using this and other bacterial strains, have not found any deoxypurines. The quantity of material reported was sufficiently great (0.2 $\mu M$ mg dry weight of cells) that it would seem probable that this material would have been observed by others.]

---

[1] In liver, following removal of a portion of the organ, tissue regeneration occurs. When DNA synthesis is followed during this process no synthesis is observed until 24-36 hours, at which time a striking synchronous increase in DNA content is observed. For this reason, regenerating liver presents an ideal biochemical system for studying certain problems in the control of DNA biosynthesis (Barnum *et al.*, 1957).

As with regenerating liver, bacteria do not exhibit a gross build up of deoxynucleotides prior to synchronous DNA synthesis nor are purine deoxyriboside derivatives accumulated in even measurable quantities during any portion of the cell cycle (Lark, 1961). Moreover, those slight changes in the deoxynucleotide pool which are observed have been shown to be unrelated to DNA synthesis since they may be retained as discontinuous changes under conditions leading to continuous DNA synthesis (see Fig. 6). In addition, if a step of DNA synthesis is blocked, no measurable excess of deoxynucleotides is accumulated.

Again, the sudden onset of DNA synthesis in *Salmonella typhimurium,* which results from raising the temperature from 25° to 37°C,

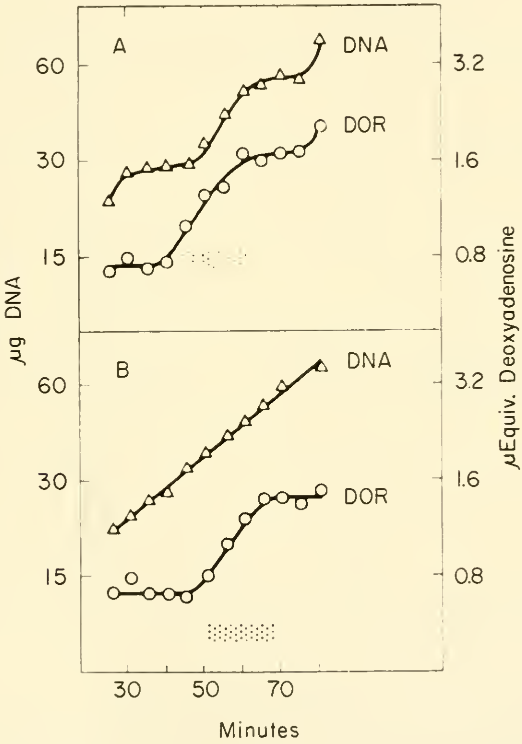


FIG. 6. Acid-soluble deoxyribotide (DOR) synthesis in relation to discontinuous DNA synthesis in synchronized bacteria. Deoxyribotides are given as milligram equivalents. Deoxyribotides are synthesized discontinuously about (5 minutes) before the onset of DNA synthesis (A). The quantity of material synthesized is less than 3% of the intracellular content of DNA. If the discontinuity of DNA synthesis is removed (B) by the addition of deoxyribosides to the medium during a previous cycle of synthesis (see text, p. 196), the pattern of subsequent deoxyribotide synthesis is unaltered. (From Lark, 1961.)

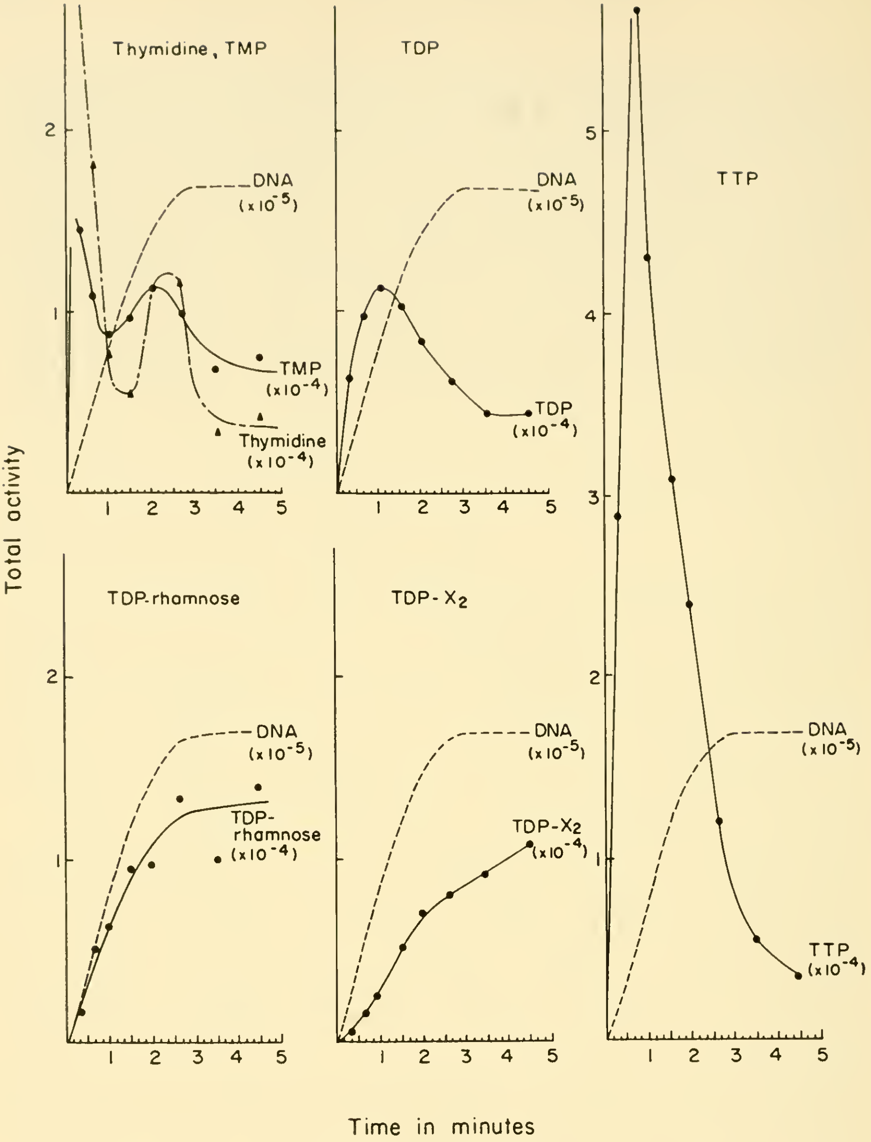


FIG. 7. The pattern of acid-soluble thymidine utilization by *E. coli* B. A trace of tritiated thymidine was added to the growth medium of the bacteria, and was rapidly utilized. Cells taken at the intervals indicated were analyzed for their intracellular acid-soluble thymidine. The distribution in each of several components is given. As may be seen, those components which have been shown to represent the major steady state pool constituents, TDP rhamnose, and TDP-$X_2$ (both thymidine diphospho sugars) do not appear to serve as immediate DNA precursors. The minor pool constituents, TMP, TDP, or TTP, appear to turn over (incorporate and lose radioactivity) rapidly as would be expected. (Unpublished data of Dr. R. Okazaki presented here by his kind permission.)

is not preceded by a rapid increase in intracellular deoxyribosidic compounds (Lark, 1961).

Okazaki has shown that the major components of the deoxynucleotide pool in *E. coli* do not normally serve as an immediate precursor for the bulk of the intracellular DNA. Thus, labeled thymidine incorporated into thymidine diphospho sugar compounds does not turn over as rapidly as thymidine mono-, di-, or triphosphates (Fig. 7). Studies on *Lactobacillus acidophilus* have shown that such thymidine compounds *can* eventually serve as a source of material for DNA synthesis; however, the rate of transfer of label is quite slow (Okazaki *et al.*, 1959). Experiments of this type are extremely valuable in distinguishing whether a compound found in a cell is, in fact, serving as an immediate precursor for DNA synthesis.

Feinendigen *et al.* (1961) have obtained a similar result to that of Okazaki by studying the rate of incorporation of cytidine and of thymidine into DNA of animal cells. They used autoradiographic techniques and demonstrated that cytidine entered a pool from which it could be withdrawn over a long period for use in DNA synthesis. This pool could not be affected by exogenously supplied material. In contrast, thymidine entered DNA via a much smaller pool, susceptible to changes in the outside medium.

With the exception of the lily anther system and developing amphibian or echinoderm embryos, the majority of data strongly support a picture in which deoxynucleoside or deoxynucleotide precursors are supplied for DNA synthesis upon demand, neither accumulating to any great extent before synthesis nor diminishing greatly during synthesis. This conclusion is supported by autoradiographic experiments which have demonstrated that the pool which can dilute exogenous thymidine or deoxycytidine during incorporation into DNA is extremely small (Painter *et al.*, 1960; Taylor, 1960b).

Little is known about the control of DNA synthesis resulting from the conversion or lack of conversion of ribotides to deoxyribotides. However, *in vitro* studies have indicated that a singular type of feedback inhibition mechanism may exist (Reichard, 1960; Reichard *et al.*, 1961). Purine or pyrimidine deoxyriboside triphosphates have been found to inhibit the *in vitro* enzymatic conversion of nucleotide monophosphates to deoxynucleotide phosphates. Moreover, this inhibition is not specific for the particular type of purine involved, i.e., deoxyadenosine triphosphate will inhibit the conversion of cytidylic acid to deoxycytidine monophosphate. Recent studies have shown that this system may be operative *in vivo* as well as *in vitro*. Several workers have found that high concentrations of deoxynucleosides in the medium will inhibit the synthesis of DNA *in vivo* in both bacteria and animal cells (Prusoff, 1959; Morris and Fischer, 1960; Lark, 1960; Longer and Klenow, 1960). Longer and Klenow (1960) have shown that whereas deoxyadenosine will inhibit the synthesis of DNA by ascites tumor cells, this inhibition can be overcome by the addition of deoxyguanosine. The inhibition is accompanied by the accumulation, intracellularly, of deoxyadenosine triphosphate (Munsch-Peterson, 1960). Other studies have indicated that this is responsible for the block in DNA synthesis (Overgaard-Hansen and Klenow, 1961; Klenow, 1961). On the other hand, these effects have been demonstrated to date only in the presence of extremely high concentrations of deoxynucleosides in the medium, whereas purine deoxyriboside triphosphates have not been observed to accumulate in normally growing cells. For these reasons it is not clear whether this method of regulation is used to control the supply of precursors needed for DNA biosynthesis.

The possibility also exists that the rate-limiting step in the synthesis of DNA precursors may be the availability of ribonucleotides from which deoxyribonucleotides are formed. A system of this type has been obtained from phage-infected cells by Cohen *et al.* (1961), who have studied the *in vitro* conversion of RNA to deoxyribonucleotides. They have noted the fact that RNA may play a key role in the control of DNA and protein biosynthesis (see Section III below) in that an RNA molecule may serve as substrate in the synthesis of the former or template in the synthesis of the latter but cannot be both template and substrate at the same time. They point out, however, that such a control mechanism can only operate in systems where ribonucleotides are limited in availability. This does not appear to be the case in bacteria. In synchronized bacteria, in which DNA is being synthesized in steps, large amounts of ribonucleotides are present at all times during the cell cycle, (Maruyama and Lark, 1962) and bacteria, which synthesize DNA and RNA at different rates, adjust the intracellular concentration of ribonucleotides to correspond to these rates (Franzen and Binkley, 1961; Smith and Maaløe, 1962). In all cases, the intracellular ribonucleotide content is ample to provide material for DNA synthesis, and DNA and protein synthesis proceed simultaneously.

### B. THE MEASUREMENT OF ENZYME ACTIVITY ASSOCIATED WITH THE PRODUCTION OF DNA PRECURSOR

This indirect approach has been used by several workers interested in the enzymatic basis of the control of DNA synthesis.

In interpreting data of this type, it must be remembered that the ability to demonstrate active enzyme in extracts of a tissue does not mean that the same enzyme is necessarily active *in vivo*. On the other hand, the absence of a necessary enzyme may indicate a negative type of control over DNA synthesis (provided the extraction procedure would preserve the enzyme activity).

The isolation and identification of the enzymes involved in the production of DNA precursors have been carried out using a wide variety of biological sources (see Chapter I). As was pointed out for the case of bacteriophage, it has been possible to demonstrate the *de novo* synthesis of various enzymes necessary for the production of specific types of DNA.

Few attempts have been made, however, to correlate pertinent alterations in intracellular enzyme patterns with the DNA cycle. We have already noted the induction of phosphorylating enzymes in the microspores of the lily anther (Hotta and Stern, 1961b). The time sequence was such that it was suggested that activity had been induced by the deoxyriboside substrates. A more detailed study has been made of a similar situation in regenerating liver.

The liver of adult rats contains all of the enzymes necessary to convert deoxyadenylic acid, deoxycytidylic acid or deoxyguanylic acid to the corresponding triphosphates (Hecht *et al.*, 1954; Hecht and Potter, 1956; Canellakis *et al.*, 1959; Gray *et al.*, 1960). However, enzymes necessary to accomplish the phosphorylation of thymidine are lacking (Bollum and Potter, 1959; Canellakis and Montsavinos, 1958; Canellakis *et al.*, 1959; Gray *et al.*, 1960; Hiatt and Bojarski, 1960). The latter appear to be present in newborn animals and are lost as the animals grow older (Hiatt and Bojarski, 1960). As a result, liver slices, as well as extracts of liver from normal adult animals, cannot synthesize DNA *in vitro*.

Following partial hepatectomy the necessary thymidylic kinases can be isolated from the regenerating liver (Bollum and Potter, 1959; Canellakis *et al.*, 1959; Gray *et al.*, 1960; Hiatt and Bojorski, 1960; Bianchi *et al.*, 1961) at the time at which DNA synthesis commences.

It is natural to assume that it is the production of these enzymes and the consequent production of the required thymidine triphosphates that leads to resumption of DNA synthesis. In examining the question of the control of DNA biosynthesis, however, several questions arise:

*a*. Does control lie in the production of a single enzyme or of a sequence of enzymes? Recent evidence indicates that the enzymes for thymidine production may not be produced until regeneration commences (Hiatt and Bojarski, 1960). Moreover, a careful examination of the time of appearance of thymidine, thymidine monophosphate, and thymidine diphosphate kinases has shown that these enzymes appear sequentially in that order during regeneration (Weissman *et al.*, 1960). Similar results were obtained by the same workers studying enzyme production by L cells of mouse fibroblasts grown in tissue culture. In this case, the sequential appearance of enzymes was observed upon transfer of cells to fresh medium. Finally, it has been shown that production of these enzymes by normal liver can be induced by injecting animals with thymidine (Hiatt and Bojarski, 1960). Thus, it would appear that a sequence of enzymatic steps, or rather a metabolic pathway, is concerned and that this pathway may be normally repressed or inhibited. In this connection, it is interesting to note that normal liver contains an extractable substance which will specifically inhibit the thymidylic kinase system (Gray *et al.*, 1960).

*b.* Is the establishment of this pathway the sole controlling factor in DNA biosynthesis? If so, one would expect that the maximum rate of DNA synthesis would not be achieved until the entire enzymatic mechanism for the synthesis of thymidine triphosphate was optimally established. Instead, evidence exists which indicates that the maximum enzyme activity (both thymidylic kinase and polymerase) is reached after the rate of DNA synthesis has already begun to decline (Bollum and Potter, 1959; Bianchi *et al.*, 1961). Similarly, when DNA synthesis ceases in regenerating tissue, precursors continue to accumulate (Williamson and Gaschlbauer, 1961). Thus, if synthesis of thymidine triphosphate is responsible for the onset of DNA synthesis, the latter does not cease as a result of an absence of the former.

Canellakis *et al.* (1959) observed the presence of enzymes in normal liver, which degrade uridine or thymidine compounds. A reduction in their activity in regenerating liver was thought to contribute to the control of DNA synthesis. In the same system an increase in thymidylate kinase activity was observed. However, the observed changes in activities of all of these enzymes were not cyclic, whereas DNA synthesis took place in three cycles occurring at 24, 48, and 72 hours. Clearly, if only the production of these enzymes controlled DNA biosynthesis their activity should also have varied in a cyclic manner.

From these considerations it would appear that although the production of enzymes needed for the synthesis of DNA precursors may be necessary, it is not a sufficient condition to explain the degree of control observed.

*c.* Are other control mechanisms also operating? Evidence from X-radiation studies has indicated that it is possible to block thymidylate kinase synthesis by irradiating regenerating liver at 6 hours whereas this is impossible at 16 hours. In the latter case, however, DNA synthesis is blocked (Bollum *et al.,* 1960). Thus, it appears that with respect to radiosensitivity, at least two possible control steps may be differentiated, one of which does not involve the enzymatic production of precursors. It is not known whether both are operational in non-irradiated cells.

## III. Role of Protein or RNA Synthesis in Controlling DNA Synthesis

The relative synthesis of RNA, DNA, and protein within growing cells has been studied with increasing intensity in recent years. Such studies have stressed the importance of the interrelationship between the synthesis of different types of macromolecules. Several studies have been carried out attempting to elucidate the role that protein or RNA synthesis may play in the synthesis of DNA.

It is obvious that protein synthesis must be maintained in dividing cells if each cellular unit is to receive a sufficient complement of enzymes to function properly. Absence of certain enzymes, as we have seen in the preceding section, may help to account for the lack of DNA synthesis in certain situations. This may be the case in systems such as regenerating liver and the microspores of the lily anther (discussed above), which are unable to synthesize DNA and may be induced to regain this capacity. Although experiments with such systems are illustrative of a possible control mechanism, it is questionable whether cells which are *rapidly* dividing and synthesizing DNA lose the enzymes controlling DNA synthesis during the non-synthetic portion of the DNA cycle.

Measurements of the persistence of such enzyme systems in bacteria have shown their presence, in adequate amounts, in cells in which protein or DNA synthesis, as well as cell division, have been blocked for periods greater than a division cycle (Billen, 1960a; Doudney and Billen, 1961; Billen, 1962a). In a synchronized bacterial system, exhibiting periodic DNA synthesis, enzymes which convert thymidine monophosphate into thymidine triphosphate double in activity during a defined, short period of the division cycle. Although this increase in activity does not occur in the presence of chloramphenicol, cyclic DNA synthesis persists and enzyme activity is maintained (Maruyama and Lark, unpublished data).

In contrast to these findings, work with plants and animal cells has shown that enzymatic activity may disappear during a period of several days (or a few generations) and that some enzymes may be lost more easily than others. Thus, upon transfer of HeLa or L cells to fresh medium, the intracellular activities of thymidine kinase, thymidylic kinase, and thymidine diphosphate kinase increase sequentially (Weissman *et al.*, 1960). However, as cell growth progresses, the enzyme activities are again reduced, thymidine diphosphokinase being lost more rapidly (80% of its activity disappearing in 24 hours). Addition of thymidine to the medium appears to stabilize the enzymes (Weissman *et al.*, 1960; Bojarski and Hiatt, 1960). A similar sequence of events has been observed in regenerating liver. Such loss of enzymatic activity does not appear to be unique to the nucleotide kinases since similar findings have been observed with glutamyl transferase (Fottrell and Paul, 1961). Similarly, phosphorylating activity which has appeared in the microspores of *Trillium* shortly before DNA synthesis also disappears again within a short time (Hotta and Stern, 1961b). Thus, loss of enzymatic activity during the course of the DNA cycle could occur.[2] However, in no case has the loss of activity been sufficiently rapid to account for cessation of DNA synthesis within 4–10 hours (by the end of the synthetic period) nor can the observed dependence of enzyme stability on the medium composition and/or the conditions of growth be reconciled with a precise control mechanism which halts DNA synthesis after an exact duplication of the intracellular DNA content. Thus, although enzyme biosynthesis appears to exercise a negative type of control over the initiation of DNA biosynthesis, as in the case of regenerating liver, such a mechanism cannot account for the degree of control observed.

A more important influence in the regularly dividing cell may be the synthesis of proteins or histones necessary to stabilize the structure of the new DNA. It is quite possible that DNA synthesis is linked to the formation of such proteins by a mechanism whereby synthesis of the former cannot proceed without simultaneous or prior synthesis of the latter (see Section IV).

To answer these questions, a more general approach to the role of protein synthesis in regulating DNA formation has been devised by several workers who have studied the effects of inhibiting protein synthesis on subsequent DNA synthesis. Two types of inhibition have been employed: (*a*) starvation for an essential amino acid, and (*b*) addition of an agent which inhibits protein synthesis.

It has been generally observed that bacteria starved for amino acids required for growth synthesize a small amount of DNA, 50–150% of the pre-existing amount, after which all further synthesis ceases (Pardee and Prestidge, 1956; Barner and Cohen, 1957; Goldstein *et al.*, 1959; Okazaki and Okazaki, 1959; Gros and Gros, 1958; Billen, 1961; Maaløe and Hanawalt, 1961). In such starved cells, all demonstrable protein and RNA synthesis ceases. From these observations it was concluded that synthesis of protein and/or RNA was required for sustained DNA synthesis (Goldstein *et al.*, 1959; Maaløe and Hanawalt, 1961). More careful investigation of the synthesis of DNA in cells starved for an essential amino acid revealed that it could proceed to a limited extent in individual cells before ceasing. (In these studies the cells were also starved of uracil; however, the similarity of the behavior of the culture as a whole to experiments of others in which only amino acids were limited makes it most probable that the effects observed were the direct result of amino acid starvation.) The amounts of DNA made (ca. 50% of the initial cell content), as well as the number of cells which at any time during starvation could be shown (by autoradiography) to retain the capacity for DNA synthesis, were consistent with the hypothesis that each cell continued to finish the DNA replication cycle upon which it had embarked (Hanawalt *et al.*, 1961). Experiments in which DNA formed under these conditions is labeled by means of a density marker are also consistent with this hypothesis (Hanawalt *et al.*, 1961; Nakada and Ryan, 1961). It was suggested that such cells could not commence a second cycle of DNA replication without a short respite from the starvation to which they were subject. It was concluded that protein and/or RNA synthesis were required to initiate a new cycle of DNA replication. The evidence presented appears conclusive in showing that DNA synthesis can proceed at least to a limited extent in the absence of RNA and protein synthesis. Moreover, the hypothesis that a cell engaged in DNA synthesis will finish a cycle of replication despite the absence of RNA or protein synthesis is extremely attractive.

---

[2] Recently the possibility has been raised that the observed lack of enzyme activity in systems such as regenerating liver may be due to the conversion of intracellular enzyme from a soluble into a particulate form (Bianchi *et al.*, 1961).

Subsequent experiments have shown, however, that the amount of DNA made will depend upon the growth conditions prior to starvation and that, under certain conditions, cells can more than double their content of DNA or carry on synthesis beyond the end of one cycle. Moreover, quantitative autoradiography indicates that those bacteria which are able to retain their ability to synthesize DNA for a long time may have a relatively slow rate of DNA synthesis (Schaechter, 1961). Thus, an alternative possibility is that all of the cells retain the capacity to synthesize a certain amount of DNA upon deprivation of an essential amino acid but some cells exhaust this capacity quickly, whereas others take a longer time.

Another possibility, which has not been properly resolved, is that under certain starvation conditions bacteria may complete one replication cycle in addition to that in which they were engaged when starvation began (possibly by using an endogenous reservoir of the essential amino acid).

The bacteria investigated have been shown to synthesize DNA throughout the majority of their division cycle. When a similar experiment was carried out on animal cells grown in tissue culture, which have a defined period, *S*, of DNA synthesis, they were found, after 48 hours of amino acid deprivation, to be arrested in a state in which the DNA content of the nucleus was that characteristic of cells which have just divided (Paul, 1962) (see Fig. 8).

Similarly, when an amino acid analog, ethionine, was fed to rats, it was found that the DNA content of individual liver cells was one-half that ($4C \rightarrow 2C$) found in control animals or in animals in which the inhibition was overcome by feeding methionine (Holzner *et al.*, 1959). The results of this latter experiment are similar to the finding of Schneider *et al.* (1960) that injections of certain amino acid analogs will result in a loss of ability to incorporate thymidine-$H^3$ into regenerating liver. This may be due to inability to form active DNA-synthesizing enzymes, or possibly to an effect similar to that observed in bacteria starved for amino acids. Unfortunately, no enzyme assays were carried out and thus it is impossible to distinguish beween the two hypotheses (Schneider *et al.*, 1960). Similar observations to these were made by Lieberman and Ove (1962), who studied DNA synthesis in cells from rabbit kidney cortex.

Further experiments would seem to be required in the animal cell systems in order to determine at what stage in $G_1$ such cells are halted; whether or not individual cells are all halted in the same stage; and the number of DNA cycles which can take place following the onset of starvation. However, these findings would tend to support the hypothesis that amino acid deprivation allowed cells to finish DNA duplication and division but prevented initiation of a new cycle of DNA biosynthesis.

The assumption that protein synthesis is required for the initiation of a new DNA synthetic cycle needs to be examined in greater detail. Although the effects of chloramphenicol on *E. coli* (Billen, 1959b; Neidhardt and Gros, 1957; Doudney, 1961a) tend to support the observations on cells starved for amino acids, frequently the magnitude of the chloramphenicol effect does not parallel that of amino acid starvation. This is exemplified by experiments with *Lactobacillus acidophilus* (Okazaki and Okazaki, 1959) for which it has been shown that deletion of essential amino acids from the medium results in suppression of DNA synthesis following a pattern observed in the starved auxotrophs of *E. coli* described above (residual synthesis of 80–100% of existing intracellular content). However, if RNA and protein synthesis were inhibited by a combination of uracil starvation and the addition of chloramphenicol, no inhibition of DNA synthesis was observed. A similar inhibition of RNA

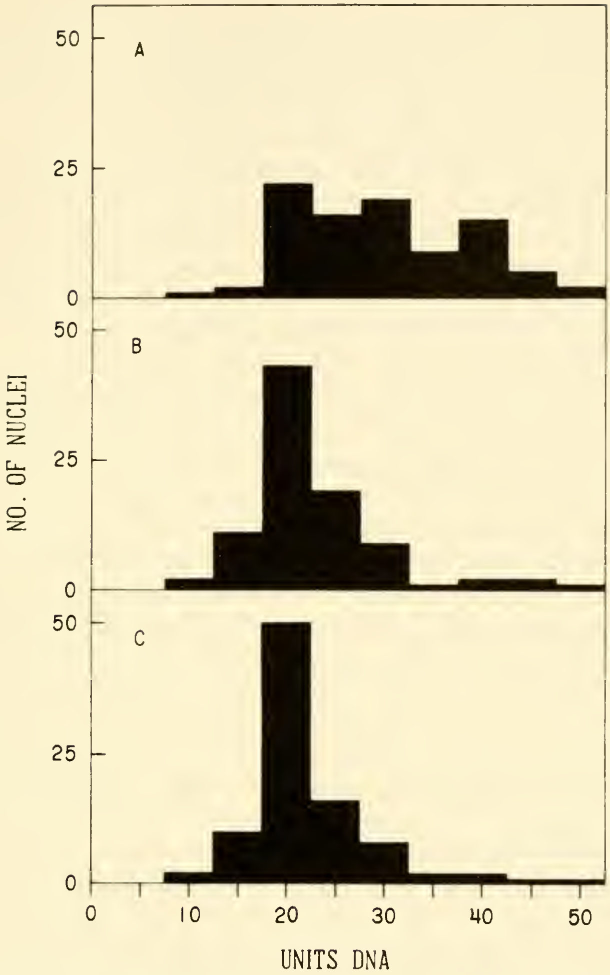


FIG. 8. Microspectrophotometric distribution of the DNA content of individual L cells (clone P2) grown for 48 hours in (A) complete medium, (B) medium from which leucine was omitted, and (C) medium from which glutamine was omitted (10 units of DNA equal to approximately 4 $\mu\mu$g). In the absence of either leucine or glutamine nuclei accumulate which cannot double their DNA complement. (Unpublished data of Dr. J. Paul printed here by his kind permission.)

and protein synthesis could be achieved by simply prolonging the period of uracil starvation, and again no diminution of DNA synthesis was observed (in both cases synthesis of DNA was greater than 600% of the initial intracellular content). As an example of a possible type of explanation which could be advanced, the authors pointed out that amino acid-conjugated deoxynucleotides might be required for DNA synthesis.

It has also been found that in synchronous cultures of bacteria in which the DNA synthetic cycle consists of a step of synthesis followed by a period of no synthesis, chloramphenicol blocks synthesis of cytoplasmic protein but will allow DNA synthesis to continue until the intracellular content of the latter is more than doubled. Moreover, DNA continues to be synthesized (to a reduced extent) in steps separated by periods equal to those found in non-treated cultures (Maruyama and Lark, 1961). Periodic DNA synthesis was followed under these conditions for three cycles (during which the cellular DNA content doubled. This would argue that protein synthesis did not control the initiation of (or the timing of) a new step of DNA synthesis.

In these experiments, the rate of DNA synthesis (that is, the amount made in a definite time interval) was found to depend upon the time at which chloramphenicol was added to the culture. Thus, it appeared that the rate of DNA synthesis was increased if protein synthesis was allowed to proceed during the initial stages of the cell cycle.

A similar study was made on bacteria with synchronous division cycles in which no cyclic synthesis of DNA occurred (Doudney, 1960). In this system DNA synthesis was inhibited only if chloramphenicol was added at a certain stage of the cell cycle. The author concluded that DNA synthesis depended upon protein synthesis occurring during that particular phase of the cell cycle. This conclusion, however, appears unwarranted since cells subjected to chloramphenicol 10 minutes earlier were perfectly capable of synthesizing DNA for the next 40 minutes, that is, during and after the period in which it was postulated that protein synthesis was required. This occurred despite the fact that protein synthesis was blocked.

An extension of experiments on the starvation of amino acid auxotrophs has been carried out by Billen (1960b), who has shown that the recovery of the ability to synthesize DNA upon subsequent refeeding can be inhibited by irradiation or by chloramphenicol. As observed by others, chloramphenicol was not found to be as effective as amino acid starvation in blocking DNA synthesis. However, it blocked resumption of DNA synthesis previously inhibited by amino acid starvation. These experiments suggest that during the process of amino acid starvation proteins necessary for DNA synthesis are lost or destroyed and that chloramphenicol or irradiation may destroy or inhibit the mechanism for replacing these.[3] Experiments have shown that cells thus treated retain all of the active enzymes necessary for DNA synthesis, i.e., all of the kinases and the polymerase (Billen, 1960b, 1962a). Therefore, if a critical protein is lost or destroyed, it would have to be of a type other than those involved in the enzymatic synthesis of DNA. Cells in which DNA synthesis has been delayed as a result of irradiation or nitrogen mustard treatment also do not resume this activity if protein synthesis is inhibited (Harold and Ziporin, 1958; Billen, 1959a,b; Harold and Ziporin, 1959; Draculic and Errera, 1959) and, again, either chloramphenicol or starvation of an essential amino acid will produce similar effects (Billen, 1961). Extracts of such irradiated cells also do not appear to have lost their enzymatic capacity for *in vitro* DNA synthesis (Doudney and Billen, 1961).

---

[3] This might occur in a manner similar to that in which sulfur-starved cells lose alcohol-soluble proteins (Roberts *et al.*, 1955).

Thus, although it appears probable that protein synthesis is normally not necessary to allow a cycle of DNA synthesis once initiated to run to completion, further experimentation is required before it is possible to conclude that protein synthesis must occur in order to initiate a new cycle of synthesis. There can be no doubt that deprivation of amino acids from a cell which requires them will inhibit DNA synthesis. However, it has recently been shown (Gallant and Suskind, 1962; Kellenberger *et al.*, 1962) that this inhibition can be reversed by the addition of chloramphenicol alone. As the period of starvation is increased, however, it becomes more difficult to reinstate DNA synthesis either by addition of chloramphenicol or by readdition of amino acids. This second, irreversible effect may represent a true requirement for protein synthesis.

Amino acid starvation has also been shown to block RNA synthesis. Moreover, evidence has accumulated which indicates that only certain types of RNA may accumulate in the presence of low concentrations of chloramphenicol and that such RNA may be unstable (Neidhardt and Gros, 1957; Astrachan and Volkin, 1958; Britten *et al.*, 1962; Nomura and Watson, 1959). Further investigation (Aronson and Spiegelman, 1961a,b) has shown that the amount of RNA synthesized, its stability, and perhaps the type of RNA will depend on the concentration of chloramphenicol used and may depend on the amino acid composition of the medium. Thus, under certain conditions, RNA will accumulate in chloramphenicol-treated cells, whereas in others it may be broken down. Radiation will also inhibit the synthesis of RNA as well as of protein. Thus, in several of the experiments cited above, the controlling factor may be the synthesis of a particular type of RNA rather than protein.

This is illustrated by several experiments.

It has been shown that the prolonged synthesis of protein and RNA, under conditions in which DNA synthesis is prohibited, results in an increase in the amount of DNA which can be subsequently made in the absence of protein and RNA synthesis. This has been done with a mutant of *E. coli* B, T⁻ hist⁻, requiring thymine and histidine, starved of thymine in the presence of histidine, and subsequently starved of histidine in the presence of thymine. From the data it appears that the amount of *new* DNA synthesized appears to be proportional to the net amount of RNA which can be synthesized. Chloramphenicol could be shown to block this increased capacity for DNA synthesis (Nakada, 1960).[4]

Bacteria starved of an amino acid (tryptophan) are unable to synthesize DNA in the presence of chloramphenicol and tryptophan. Upon preincubation with amino acid, starved cells recover the ability to synthesize DNA in the presence of the drug. The rate at which this ability is recovered has been correlated with the synthesis of RNA (Doudney, 1961b). A similar result was obtained if DNA synthesis was blocked as a result of UV irradiation rather than amino acid starvation (Doudney, 1961a). In both of these experiments some correlation could, of course, be made with the recovery of protein synthesis which is known to depend on RNA synthesis. The frequency of ultraviolet induced mutations (shown to be maximal if DNA synthesis is allowed to proceed as rapidly as possible) also appears to be greatest if the synthesis of RNA is not inhibited (Doudney and Haas, 1961).

Finally, it has been demonstrated that the genetic inability to synthesize RNA in the absence of amino acids (Stent and Brenner, 1961) also extends to the synthesis of DNA (Kellenberger *et al.*, 1962).

Although these findings imply that RNA rather than protein synthesis may be necessary for the initiation of DNA synthesis, the results of Okazaki and Okazaki (1959) argue that this is not the case, since they obtained DNA synthesis without RNA synthesis when lactobacilli were grown in the absence of uracil. It has been suggested (Kellenberger *et al.*, 1962) that both RNA and DNA synthesis are subject to the control mechanism involving amino acid participation.

## IV. The Physical State of DNA as a Factor Controlling Its Replication

The requirement for pre-existent DNA to act both as a template and a primer in the synthesis of DNA has been well established by *in vitro* studies (see Chapter I). The availability of intracellular DNA to act in this capacity may play an important role in controlling its own biosynthesis. This has been suggested by the following experiments:

---

[4] Although the investigator noted a correlation between the amount of protein synthesized and the amount of DNA which can be made, this appears to be erroneous, since the comparison is made between net contents of DNA rather than newly synthesized DNA. From the data, the correlation with newly synthesized RNA appears to be more significant.

Cultures of synchronized bacteria (*Alcaligenes faecalis*), in which DNA synthesis occurs in steps, will cease to synthesize DNA if large amounts of deoxyadenosine are added to the medium (presumably as a result of precursor starvation; see Section II,A). If this is done immediately before a burst of DNA synthesis (to cells which presumably have completed all the necessary preparations for DNA synthesis) DNA synthesis is blocked. Reversal of inhibition at different periods thereafter leads to an eventual resumption of DNA synthesis which occurs during a period in which a control (non-inhibited) culture undergoes a second cycle of DNA synthesis. The timing at which this eventual DNA synthesis takes place is independent of when inhibition is initiated or reversed (Lark, 1960), and is illustrated in Fig. 9. Non-synchronized exponential cultures of this organism treated in this way initiate DNA synthesis immediately upon reversal of inhibition and proceed to synthesize DNA exponentially at the normal rate. These results have been interpreted as evidence for the existence of certain allowed periods of DNA synthesis. The lack of correspondence between these periods and the periods for which DNA precursor production is inhibited led to the suggestion that control of these periods depended upon the conversion of DNA to a primer state rather than upon the production of precursor material. This was further emphasized by the finding that DNA precursors, when added to the growth medium, could only influence the DNA cycle if added during an allowed period of synthesis.

Cultures of *E. coli* starved of an essential amino acid cease production of DNA. Ability to synthesize this macromolecule is regained upon readdition of the essential amino acid. Reacquisition of this synthetic ability is inhibited by both chloramphenicol and irradiation. Extracts of cells so inhibited exhibit adequate activity of all of the enzymes necessary to synthesize DNA (Billen, 1962b). The lack of *in vivo* synthesis has been interpreted as an inability of intracellular DNA to enter into the primer state.

The slime mold *Physarum polycephalum* undergoes synchronized mitosis. Guttes and Guttes (1961) studied the ability of a "well-fed" plasmodium to support DNA synthesis in an implanted plasmodium which had been previously starved. DNA synthesis in the former took place 3–4 hours after mitosis, whereas in starved plasmodia it occurs after 6–9 hours. If a starving plasmodium was implanted into one which was well fed at a period in which neither was synthesizing DNA, the well-fed plasmodium could initiate and support DNA synthesis of the implant despite the fact that it was not itself synthesizing DNA. They interpreted their results to indicate that DNA synthesis in the well-fed plant was controlled by some factor other than the substrate and enzyme conditions required for DNA synthesis and suggested that the ability of

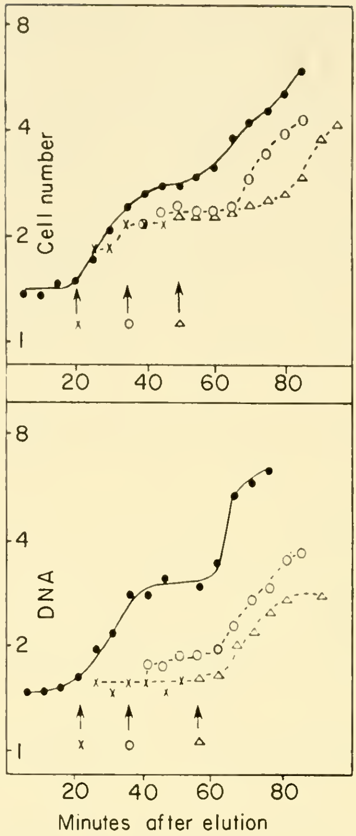


FIG. 9. Pattern of DNA synthesis and cell division following removal of the inhibitor deoxyadenosine (DOA). ●——●, control; ×——×, DOA; ○——○, no DOA after 35 minutes; △——△, no DOA after 50 minutes. DOA was added at a high concentration ($10^{-2}\,M$) immediately before the onset of DNA synthesis. Inhibition was reversed at different times (as indicated) by the addition of adenosine. Resumption of DNA synthesis occurred at the same time as the second step of DNA synthesis in the control culture despite reversal of inhibition at different times. In contrast, inhibition of cell division continued in proportion to the length of time DOA was present. The timing of DNA synthesis appears to be independent of the DNA content of the cell, or of whether or not a previous cycle of DNA synthesis was completed. (From Lark, 1960.)

DNA to act as primer was the controlling factor. (Reciprocal implant experiments demonstrated that starved plasmodia did not contain an inhibitor of DNA synthesis.)

In all of these studies the hypothesis of primer control was suggested because of the inadequacy of alternative hypotheses. In no case was a primer-controlled mechanism directly demonstrated.

The dependence of primer activity on the physical state of DNA may be inferred from *in vitro* studies of the primer requirements of cell-free polymerizing systems. The availability of primer for DNA duplication may involve several factors: (*a*) Is it necessary for DNA to exist in a special form in order for it to be duplicated? (*b*) Is DNA synthesis compatible with simultaneous RNA synthesis, and if not, to what extent are these two processes incompatible? (*c*) Is DNA structurally stabilized within the cell and is this a necessary part of DNA replication? If so, factors needed for this stabilization may control its replication. (*d*) Is the replication of DNA physically ordered, with respect to the macromolecule itself, with a beginning, a middle, and an end? If so, a control mechanism may involve the appropriate conditions for initiation of synthesis at one end of the molecule, maintenance of synthesis throughout the molecule, and termination of synthesis once the end of the molecule is reached. The onset and cessation of synthesis may also involve the activation of information contained within the molecule itself.

Little experimental evidence is available to answer these questions. Thus, the following discussion will be largely speculative.

### A. IS IT NECESSARY FOR DNA TO EXIST IN A SPECIAL FORM IN ORDER FOR IT TO BE DUPLICATED?

*In vitro* studies of the deoxynucleotide triphosphate polymerase have shown that enzymes isolated from a variety of sources require primer in order to act (see Chapter I). The ability of DNA to function as a primer in these reactions varies. In some systems native DNA appears to function efficiently as a primer, whereas in others heated or denatured DNA is required.

It appears difficult, therefore, to reach a general decision from such experiments as to whether the polymerizing reaction utilizes a special form of DNA as primer. Certainly, in some systems, denatured DNA appears to be required.

It has been shown that denaturation of DNA by heat involves strand separation (Meselson and Stahl, 1958). Similar separation occurs as a result of DNA duplication in that one-half of the parent DNA molecule is found together with an equal amount of new material in a "hybrid," semi-conserved, daughter molecule (Meselson and Stahl, 1958; Sueoka, 1960; Taylor *et al.*, 1957; Forro and Wertheimer, 1960). Despite these findings, the presence during the DNA cycle of a distinct molecule or species, which could serve as primer DNA, has seldom been observed. In one experiment (Cavalieri and Rosenberg, 1961c), it was possible to isolate a special species of DNA shortly before DNA replication in a culture of *E. coli* $T^-$ in which discontinuous DNA synthesis had been initiated by means of previous thymine starvation. This material was characterized by a greater susceptibility to DNase than normal DNA. The kinetics of DNase action were interpreted as evidence that the molecule contained half of the number of component strands usually present (Cavalieri and Rosenberg, 1961a). In all of its characteristics it resembled DNA obtained at other times from the same organism but heated in cesium chloride. On the basis of their findings these workers suggested that there exist two types of DNA molecules during the replication cycle: before DNA synthesis the molecule is cleaved and subsequently rebuilt during the replication process. Further studies of this type would be desirable, since no evidence was obtained to indicate whether or not the observed change in state of the DNA was reversed following DNA synthesis. Evidence was presented to show that DNA existed in a different state in non-proliferating as compared to dividing cells (Cavalieri and Rosenberg, 1961b).

Indirect evidence for a change in the state of DNA during its replication has been obtained by a study of the action spectrum for the inactivation of bacteriophages by ultraviolet light. Studies on the inactivation of viruses believed to contain a single- as opposed to double-stranded DNA molecules have demonstrated that those with single-stranded DNA have a different action spectrum from those with two-stranded DNA. A shift in the action spectrum of two-stranded DNA virus to that characteristic of a single-stranded DNA virus occurs following host infection and initiation of virus replication. As virus multiplication proceeds this spectrum shift is reversed (Setlow, 1960; Setlow and Boyce, 1960; Setlow and Setlow, 1960). These results have been interpreted as evidence that during replication viral DNA passes through a single-stranded or "denatured" stage.

A change in the physical state of intracellular bacterial DNA immediately following replication has been observed by Goldstein and Brown (1962). They found that newly synthesized DNA is more resistant to sonic disintegration than the bulk of cellular DNA. The nature of their experiment excluded a determination of whether this state was achieved prior to replication.

Recent analysis of DNA isolated at intervals during the division cycle of synchronized bacteria (*A. faecalis*) has revealed the presence of small amounts of material whose density corresponds to that of heated or denatured DNA. The amount of material observed (which is sensitive to DNase and can be heat renatured) is usually less than 10% of the total DNA content of the culture (Lark, Cavalieri, and Rosenberg, unpublished results).

Much less, or no such material, is found in cultures of randomly dividing cells. The "denatured DNA" arises when the cells are lysed with

Duponol to obtain their DNA. The temperature at which lysis is carried out and the *intra*cellular environment of the DNA at the time of lysis, appear to be important in determining the quantity of "denatured DNA" obtained. The evidence indicated that the intracellular environment is determining the temperature at which denaturation of the cellular DNA takes place. During the DNA cycle the amount of denatured material isolated varies. Further studies in this and other systems should facilitate the investigation of control mechanisms in which primer requirements are involved.

All of the evidence available indicates that in the process of replication the DNA molecule must change its state (splitting the molecule in half). However, it is not certain that this does not occur as an automatic result of the replication process instead of as a first step. From the experiments cited, it would appear that the latter may be true. Such a step may then represent an important controlling factor in the biosynthesis of DNA.

There is no evidence for the existence of a biological system capable of converting native into primer DNA. It is tempting to assign such a role to DNase and an increase in activity of this enzyme has been shown shortly before DNA synthesis occurs in regenerating liver (Brody and Balis, 1959). *In vitro* studies have shown that a brief treatment with this enzyme results in strand separation measured as a large increase in the reactivity of the nucleoside bases with formalin (Sarkar and Dounce, 1961; Sarkar, 1961). Moreover, DNA treated in this way becomes active when tested in a biological system requiring denatured DNA as primer (Sarkar and Dounce, 1961; Dounce *et al.*, 1961; Montsavinos and Canellakis, 1961). Unfortunately, it appears that during such treatment the molecule suffers several random scissions, a result which appears to be incompatible with the lack of observed breakage in the replication of the chromosome.[5] It is, of course, possible that specific DNase action (Laskowski, 1961) might cause non-random scissions compatible with the structural integrity of DNA.

A more attractive possibility is suggested by the observations on the conditions under which the "denatured DNA" is obtained from cells. *Changes in the intracellular environment of DNA may lower the temperature of denaturation into the physiological range. When this occurs, a portion of the DNA molecule would suddenly denature and synthesis would be initiated.* Such a hypothesis could easily explain the timing of DNA synthesis and would account for the observed effects of temperature upon the DNA cycle (see Section V,C below).

*In vitro* studies would indicate that such an idea is feasible. Hamaguchi and Geiduschek (1962) have reported the denaturation of DNA at temperatures as low as 25° in the presence of certain anions and Keir *et al.* (1962) have noted that DNA can serve as primer at slightly alkaline pH's (8–9) without having been heated.

---

[5] DNA, in which scissions have occurred as a result of X-radiation *in vivo,* undergoes degradation upon replication (Kas and Draculic, 1961).

### B. IS DNA SYNTHESIS COMPATIBLE WITH SIMULTANEOUS RNA SYNTHESIS?

Is DNA, which is serving as a functional template for the formation of cellular proteins, simultaneously able to act as a template for its own synthesis? Much recent evidence has accumulated to show that in the synthesis of proteins, RNA serves as a messenger, to carry the necessary information between the DNA template and the ribosome which serves as the site of protein synthesis (Brenner *et al.,* 1961; Gros *et al.,* 1961).

Moreover, it has been shown both *in vivo* and *in vitro* that this RNA is formed in combination with DNA and will recomplex with it (Spiegelman *et al.,* 1961; Geiduschek *et al.,* 1961; Hall and Spiegelman, 1961). In one instance an RNA-DNA-protein complex thus formed was quite stable (Bonner *et al.,* 1961).

Under some conditions the ability to replicate DNA may be related to the ability to synthesize messenger RNA. Bacteria which require thymine, arginine, and uracil for growth cease to synthesize DNA when starved of arginine and uracil in the presence of excess thymine. When arginine and uracil are returned to the medium, DNA synthesis resumes gradually after a characteristic lag. This lag in ability to synthesize DNA is independent of the actual synthesis of DNA since it occurs irrespective of whether or not thymine has been omitted from the medium. An identical lag in the ability to induce β-galactosidase formation is also observed (Roberts, 1960), which is also independent of the actual occurrence of DNA synthesis. (On the other hand, gross protein and RNA synthesis are reinitiated instantly with no apparent lag.) It is possible, therefore, that DNA in such cells is in a physical state which renders a portion of it unable to serve as a template in either RNA or DNA synthesis, or alternatively, that its activity is non-specifically repressed for a short period of time.

Experiments on DNA and RNA synthesis in amino acid-starved auxotrophs of *E. coli* have indicated that synthesis of both can be regulated by a common, amino acid-dependent, control mechanism (Kellenberger *et al.,* 1962). This mechanism would appear to operate on the primer rather than the precursor level.

No *in vitro* experiments have been reported, however, in which it was sought to determine whether the synthesis of "messenger" RNA is inhibited by the simultaneous synthesis of DNA or vice versa. Some evidence has been obtained which indicates that heat-denatured DNA, which will act as primer in self-replication, is unable to function in the *in vitro* synthesis of messenger RNA (Burma *et al.*, 1961; Ochoa *et al.*, 1961; Eisenstadt *et al.*, 1962).

That DNA and RNA synthesis are mutually exclusive *in vivo* is indicated by the experiments of Sisken (1959) and of Prescott and Kimball (1961). The latter workers studied the synthesis of RNA, DNA, and protein in the macronucleus of the ciliate *Euplotes*. As mentioned previously, DNA synthesis in this case proceeds via a DNA-synthesizing band, commencing at the edge of each lobe of the U-shaped macronucleus and continuing until the two bands meet (Gall, 1959). The nucleus then divides. They demonstrated that this DNA-synthesizing band is devoid of RNA and incapable of synthesizing RNA. It is only this part of the nucleus which was shown to have these two characteristics. Whether the progression of the synthesizing band is responsible for the cessation of RNA synthesis and dissociation of RNA-DNA complexes, or the result of it, is not known. It is obvious that if the latter were true, dissociation of the RNA-DNA complex and cessation of RNA synthesis could be a controlling factor in DNA synthesis.

Such a mechanism could explain the observed dependence of DNA synthesis on the ability of cells to carry on RNA synthesis (see Section III) since RNA synthesis could be blocked in such a way as to prevent dissociation of a DNA-RNA complex, or to favor reassociation. The recent findings of Aronson and Spiegelman (1961a,b) would indicate that chloramphenicol might, under certain conditions, create a situation in which this could occur. The formation under these conditions of RNA-DNA complexes could also explain the reduction of ultraviolet-induced mutations which occurs in the absence of RNA synthesis (Witkin, 1959; Doudney and Haas, 1961), since complex formation with RNA and subsequent lack of DNA replication might help to stabilize the structure of DNA until ultraviolet-induced lesions could be repaired. On the other hand, mutations induced by a change in base structure (resulting from alkylating agents) should not be affected, as is indeed the case (Strauss and Okuba, 1960). Finally, the observed dependence of "thymineless death"[6] (Gallant and Suskind, 1961, 1962) and DNA synthesis (Kellenberger *et al.*, 1962) on conditions leading to RNA synthesis could also be explained as a stabilization of DNA (which had undergone lesions as the result of attempted synthesis in the absence of thymine) by complexing with RNA. In all of these situations, RNA synthesis may be visualized as resulting in a turnover, and therefore dissociation, of RNA-

---

[6] This occurs when bacteria which require thymine for growth are incubated in media lacking thymine.

DNA complexes with the concurrent release of the DNA molecule, which if unstable may then undergo change.

### C. IS DNA STRUCTURALLY STABILIZED WITHIN THE CELL AND IS THIS A NECESSARY PART OF DNA REPLICATION?

Much evidence has accumulated to demonstrate that DNA exists within the cell in the form of nucleoprotein (Bloch, 1958). We have already discussed the evidence which exists for the role of protein in controlling DNA synthesis. De (1961) has shown that histone is synthesized simultaneously with DNA in the chromatin of interphase plant cells. Additional protein and histone synthesis was demonstrated in the nucleolus during $G_2$, following DNA replication (see also Woodard *et al.*, 1961). In the studies cited above with *Euplotes*, it was noted that the synthesis of DNA and of protein or histone (incorporation of thymidine and of histidine) were extremely active in the same region of the synthesizing band. This region of active synthesis, however, was extremely low in its content of both DNA and protein. The region of the band immediately adjacent contained large amounts of both. It thus would appear that the synthesis of both DNA and protein proceed simultaneously. An interesting point in this connection is that this protein can apparently be synthesized in the absence of RNA, since this region of the macronucleus is devoid of this nucleic acid. Thus, this nuclear protein may be synthesized by a special mechanism which may, as pointed out earlier, be chloramphenicol-insensitive (Nathans and Lipmann, 1961). The role of this protein is unknown. However, it has been suggested that such protein may serve to hold several DNA molecules together, linking them into a chromatid (Dounce, 1959; Kellenberger, 1960). This might account for the marked instability of such large chromatids which, although they can be shown to exist within the cell, break down, upon extraction, into smaller pieces (Burgi and Hershey, 1961). The molecular weight of the latter suggests a frequent and uniform occurrence within the chromatid of weak links. It has been suggested that the synthesis of DNA may be controlled by the rate of synthesis of these "linkers" (Kellenberger, 1960).

### D. IS THE REPLICATION OF DNA PHYSICALLY ORDERED, WITH RESPECT TO THE MACROMOLECULE ITSELF, WITH A BEGINNING, A MIDDLE, AND AN END?

No evidence exists to indicate whether the DNA molecule, the chromatid, or portions of the latter are synthesized in a linear progression or not. The evidence for a preferentially timed synthesis of individual chromosomes has been discussed in a previous chapter (Chapter II). Experiments on density labeled chromatin indicate that the proportion of the chromatin in the process of being synthesized is extremely small compared to that which is either already synthesized completely or that which is not yet synthesized. Thus, only a small portion of the material is being replicated at any time (Sueoka, 1960; Meselson and Stahl, 1958). Whether this portion moves from one end of the chromatid to the other is not known. That this may be the case is indicated indirectly by the experiments of Rudner (1960) on the induction of mutations by 2-aminopurine incorporated into a synchronized culture of bacteria. In this case, appearance of mutants occurred in steps, whereas DNA was synthesized continuously. This would indicate that synthesis of particular loci in the DNA molecule occurred at precisely timed intervals.

It is difficult to see how sequential synthesis of DNA can be avoided if strand separation is to occur without frequent breakage or crossing over which would destroy the semi-conservative pattern of chromosome duplication which has often been observed. If DNA synthesis occurs in small bursts covering the rapid synthesis of one segment (molecular weight ca. $10^6$) after another of the chromatid and if these are to be stabilized by linkers of some type, it also would appear reasonable to suppose that a stabilized type of bonding (perhaps covalent) must occur at each end of the segment, to prevent subsequent rupture of terminal hydrogen bonds and possible loss of the semi-conservative pattern. Such bonding has been shown to result from ultraviolet irradiation and it is possible that it could occur in the absence of irradiation. If this were the case, initiation of synthesis could be an all-or-none event dependent upon breakage of such critical bonds (Marmur and Grossman, 1961).

A key question concerning control via a point of initiation may be: What happens to cells in which DNA synthesis is interrupted before completion? When synthesis is resumed will it resume at the point of the chromatid where it was interrupted, or at the original initiation point? We do not know whether, once initiated, the conditions necessary for DNA synthesis remain imposed upon the cell until such synthesis can be completed. To study this, it is necessary to have a system in which damage to DNA or other cell constituents does not occur during the blockage of synthesis. The use of deoxyadenosine inhibition for autoradiographic studies of chromosome synthesis in higher cells would appear extremely fruitful from this point of view.

In general, it is obvious that if DNA synthesis is initiated at one end of the molecule, special conditions will be needed to distinguish the beginning and end of the molecule and to prepare it for initiation of synthesis.

Such a statement implies that the molecule itself carries specific information which differentiates one end from the other. If this differentiation is to play a controlling role in DNA biosynthesis then the availability of this information must be under some on-off type of control, similar to the control of operator genes by cytoplasmic repressors (see Jacob and Monod, 1961). However, a similar type of cytoplasmic repressor would not appear to be involved, since conditions leading to the accumulation of RNA, DNA, or protein within the cell do not affect the timing of DNA synthesis (Maruyama and Lark, 1961), which is also independent of the synthesis or lack of synthesis of DNA itself (Lark, 1960).

Guttes and Guttes (1961) have suggested that the nucleolus may play the role of an "operator" since its synthesis is initiated simultaneously with that of DNA.

## V. OTHER FACTORS AFFECTING DNA BIOSYNTHESIS AND THE DNA CYCLE

### A. THE LOCALIZATION OF DNA SYNTHESIS

The *in vivo* synthesis of DNA depends upon its environment within the cell. Thus, fragments of DNA (resulting from radiation damage) which are left in the cytoplasm as micronuclei can no longer serve as foci for DNA synthesis although the parent nucleus still does (Das and Alfert, 1961b). In ciliates different nuclei such as the macronucleus and the micronucleus have been found to follow different patterns of DNA synthesis, the micronucleus entering its period of DNA synthesis when the macronucleus has finished and when the cell is beginning to divide (McDonald, 1962; Prescott *et al.*, 1962).

These data would indicate that the nucleus represents a unique environment for DNA synthesis; and it has been shown that specific DNA precursors such as thymidine, when taken into the cell from the environment, will rapidly appear in their free form, or as nucleotides (thymidine mono-, di-, or triphosphate), in the nucleus, but are not found in the cytoplasm (Crathorn and Shooter, 1960). However, good evidence has been accumulated demonstrating that the deoxyribonucleotide polymerase is found in the cytoplasm, but *not* in the nucleus of mammalian cells (Smellie *et al.*, 1959; Smellie and Eason, 1961; Prescott *et al.*, 1962a). In bringing together all of these components required for the synthesis of DNA, a controlling factor may be the permeability of the nuclear membrane.

A further factor may be the association within the nucleus of enzyme, substrate, and primer. Billen (1962b) has recently demonstrated that in *E. coli* the polymerase is complexed with the DNA. It is not known whether the formation of this complex is a controlling factor in DNA synthesis. *E. coli* is known to synthesize DNA throughout most of its cell cycle (see Introduction). However, the same type of complex is found at all times in synchronous cultures of *A. faecalis*, which synthesize DNA discontinuously (Billen and Lark, unpublished data).

### B. VARIATION IN THE DNA CYCLE

Whereas rapidly dividing cells in tissue culture often exhibit a precisely defined cell cycle, the same is not necessarily true for cells growing in the animal. Mendelsohn *et al.* (1960) have shown that transplanted tumor tissue shows a large variation in the length of the period preceding DNA synthesis ($G_1$), whereas both $S$ and $G_2$ remain rather constant over the cell population. Studies on the intestinal epithelium and the ear epidermis of the mouse (Quastler and Sherman, 1959; Quastler, 1960; Sherman *et al.*, 1961) have shown a similar variation in $G_1$, as opposed to $S$ or $G_2$. Moreover, different cell types from the same animal showed differences in the average length of their $G_1$ periods. Differences were also observed in $S$ and $G_2$, but these were much less striking. In addition, it was possible to show that the decision as to whether an intestinal epithelial cell should cease dividing and differentiate (move into the villi) or to continue proliferation, was taken during the early portion of the cell cycle *before* DNA was synthesized. On the other hand, cells of the epidermal layer of the ear often ceased division in the $G_2$ (Gelfant, 1960) as well as the $G_1$ (Gelfant, 1961) stages. However, the low temperature at which these cells are growing may account for a mitotic block arresting cells in $G_2$.

Upon injury, cells may move out of both $G_1$ and $G_2$ to commence DNA synthesis and/or division. This has been demonstrated with cells of the epidermal layer of the mouse ear and a similar situation may exist in tissue of the eye (Harding and Srinivason, 1961). It is, of course, known to exist in liver as evidenced by the events involved in regeneration (Barnum *et al.*, 1957). Injury would appear in these cases to provoke a humoral response capable of initiating DNA synthesis and hence cell proliferation. Thus, partial hepatectomy of one animal will result in liver proliferation in another animal which is placed in parabiosis with it (Bucher *et al.*, 1951).

It is interesting to note that during liver regeneration a good deal of the variability between the DNA cycles of individual animals (as evidenced by increasing asynchrony of DNA synthesis) may be reduced by the use of inbred strains of the same age which have been partially hepatectomized at the same time. This would indicate that the inherent rate of progression through $G_1$ may be related, in some way, to a diurnal cycle of the animal as a whole and thus be under some humoral control mechanism which does not involve the actual synthesis of DNA itself.

Although variation among individual cells is much less for animal cells growing in culture, Terasima and Tolmach (1962) have noted that most of the randomization observed at $M$ in a synchronized population of HeLa cells occurs during $G_1$. Moreover, animal cells growing in culture are halted in $G_1$ when the medium in which they are growing is exhausted (Whitfield and Rixon, 1959). The length of the DNA cycle may also change for an entire cell population when the growth conditions in culture are altered. When this occurs, the change is expressed as a lengthening or shortening of $G_1$ rather than $S$ or $G_2$ (Richards *et al.*, 1956). A similar situation may occur in bacteria where it has been shown that major increases in growth rate brought about by enriching the growth medium are accompanied by an increase in the average DNA content per nucleus (Schaechter *et al.*, 1958). This would indicate that more DNA is now being synthesized in the early stages of the cell cycle thus raising the average DNA content per nucleus.

The data available would indicate that under most conditions, $G_1$, the portion of the DNA cycle prior to synthesis, is the most susceptible to change.

### C. EFFECTS OF TEMPERATURE AND RADIATION ON THE DNA CYCLE

Although several workers have noted the effects of temperature on the division cycle, few studies have been made of the effect of heat or cold on DNA synthesis and these have yielded little information. Zeuthen and Scherbaum (1954) synchronized tetrahymena by repeatedly shifting the temperature from 29° to 34°C. At the latter temperature cell division appears to be blocked, whereas DNA synthesis continued. Subsequent experiments demonstrated that during such repeated cycles, the cells doubled their content of DNA without undergoing division (Scherbaum, 1957; Scherbaum *et al.*, 1959) but that upon being returned to their optimal growth temperature, division took place before further DNA synthesis was resumed (thus reducing the DNA content). These experiments indicate that the temperature block (in this case a heat block) was operative during the pre-mitotic, or $G_2$ period.

Randomly dividing bacteria (*S. typhimurium*) grown at 25°C have the same content of DNA as when grown at 37°, although the rate of synthesis at the former temperature is about half of that at the latter (Lark and Maaløe, 1956). However, when cells grown at 25°C are shifted to 37°, a sudden burst of DNA synthesis occurs equal to approximately 30% of the cellular content of DNA. Synthesis is then arrested for a short period. Repeated application of such temperature shifts at properly spaced intervals results in phasing of DNA synthesis and in synchronization of division. These results indicate that the population of cells growing at 25° can be divided into two classes on the basis of their ability to synthesize DNA when placed at 37° and that their DNA cycles may be divided into portions of different temperature dependence. A similar result was found with another bacterial system (*B. megatherium*) which continued to synthesize DNA at a rapid rate for a short period of time when shifted from 34° to 15° (Falcone and Szybalski, 1956). In this case, all other cell growth parameters were immediately affected by the lower temperature. Returning the temperature to 34° resulted in synchrony of cell division and phasing of DNA synthesis (Delamater, 1956). This again would indicate the existence of at least two portions of the DNA cycle with different temperature dependencies.

Hotchkiss (1954) found that the competence of pneumococci to be transformed by DNA was increased by a short period of incubation at 25°. This resulted in a synchronization of cell division and a phasing of competence. The latter, however, occurred in cycles whose frequency was somewhat different from those of cell division. Whether the cycles of competence are related to DNA synthesis, to adsorption of transforming DNA (Lerman and Tolmach, 1957), or to the ability of the transforming DNA to successfully survive within the cell until it is integrated (Fox and Hotchkiss, 1960) is not known. Further investigation of this system to ascertain the fate of the "infecting" DNA as a function of the host cell DNA cycle should be extremely profitable.

The DNA cycle of HeLa cells synchronized by cold shock differs from that of randomly dividing cells not subjected to cold (Smith *et al.*, 1959). These results indicate that two populations of cells may be created by their treatment—one whose DNA is synthesized as in cells not exposed to cold (toward the end of the cycle), and the other composed of cells in which $G_1$ has been almost eliminated and DNA is synthesized immediately after division.

The meager information on the effects of temperature on the DNA cycle indicates that different portions of the cycle may have different temperature coefficients and that the DNA cycle may be dissociated to a certain extent from the cell division cycle.

As mentioned previously, similar suggestions have resulted from studies with X-radiation. In general, exposure to X-radiation appears to block cells in $G_2$ and prevent mitosis. Cells which have not, at the time of irradiation, synthesized DNA proceed to do so before the cycle is halted although, in some cases, the onset of DNA synthesis is delayed (for reference see Seed, 1961; Whitmore *et al.*, 1961; and Quastler, 1962).

However, recent evidence has demonstrated that radiation can, under certain circumstances, initiate DNA synthesis in onion root tips before the normal onset of *S* (Das and Alfert, 1961a). Moreover, the rate of DNA synthesis early in *S* can be stimulated in both plant and animal cells by X-radiation (Alfert and Das, 1961). A similar observation was made on irradiated intestinal epithelium (Sherman and Quastler, 1960) who observed that cells in *S* matured early following irradiation. These results are consistent with the hypothesis that some chromosomes are synthesized earlier in *S* than others and that replication of unduplicated chromosomes may be initiated by irradiation. Presumably this could occur by conversion of their DNA to a primer state.

More information should become available with the study of the effects of physical changes in the environment of cells that have been synchronized by selective methods. Such studies have been initiated by Terasima and Tolmach (1962), who have studied the effect of temperature and X-radiation on cultures of HeLa cells synchronized by a selective method. Preliminary results have indicated that a period of reduced temperature (29°C) during $G_1$, or *S* will delay mitosis, whereas $G_2$ does not appear to be as temperature dependent. X-radiation during *S* will prolong this period of DNA synthesis, whereas radiation during $G_1$ is without effect. In general, radiation was found to delay or block mitosis. Most striking, perhaps is the finding that cells about to enter *S* are most sensitive to the lethal effect of X-radiation and that this sensitivity can be greatly increased by the addition of deoxyadenosine late in $G_1$ to block DNA synthesis and entry into *S* (Terasima and Tolmach, 1962).

Measurement of the rate of DNA synthesis during *S* in synchronized HeLa cells has shown that it is not constant (Terasima and Tolmach, 1962). This result confirms the indirect observations of Howard and Dewey (1961) and Painter *et al.* (1960), who came to a similar conclusion on the basis of variation in the grain count of cell labeled during random growth.

### D. THE EFFECT OF THE NUTRIENT ENVIRONMENT

Temporary deprivation of thymidine will result in synchronization of cell division and DNA synthesis in both bacteria (Barner and Cohen, 1956) and in animal cells, i.e., HeLa (Rueckert and Mueller, 1960). In the latter case, cell division ceases upon treatment with the thymidine analog fluorodeoxyuridine (FUDR), whereas attempted DNA synthesis continues for a short period resulting in the formation of a "DNA-like material." Readdition of thymidine results in resumption of DNA synthesis in two steps, just before and shortly after cell division. This result is similar to the effect of cold shock on such cells (Smith *et al.*, 1959).

These results serve to distinguish two classes of cells in the randomly dividing population. Presumably these differ with respect to when in their DNA synthetic cycle they have been subjected to the inhibitory agent.

This has been further clarified by Taylor *et al.* (1962), who have demonstrated that thymidine deprivation will only block mitosis if it occurs early in $S$. On the other hand, cells subjected to FUDR late in $S$ proceed into mitosis with a concomitant fragmentation of the chromosome. Removal of the thymidine block and addition of thymidine as late as 1 hour before anaphase will prevent fragmentation. This may indicate that incorporation of thymidine into specific sites in the nuclear material may be necessary even during $G_2$ if the integrity of the chromosome is to be maintained. This has been substantiated by Hsu and Somers (1962). They demonstrated that the number of chromosome breaks increased in proportion to the length of time prior to mitosis (that is, in $G_2$ or $S$) that cells were subjected to *low* concentrations of FUDR. The number of breaks increased continuously showing no sudden rise in number which would reciprocally represent a transition from $S$ to $G_2$.

Under certain conditions alteration of the nutritional environment may alter the DNA cycle without blocking cell growth or division. Only a few such experiments will be discussed.

Bacteria (*E. coli*) when grown exponentially under conditions of balanced growth appear to synthesize DNA throughout the majority of their division cycle (Schaechter *et al.*, 1959). However, when ribosides are added to the medium DNA synthesis becomes cyclic, occurring in steps separated by periods in which almost no synthesis takes place (Lark and Lark, 1962). No phasing was observed in cell division, although the over-all growth rate was increased. This result is related to the findings of Kjeldgaard *et al.*, (1958), who observed a delay before the rate of DNA synthesis was increased as a result of shifting bacteria to a medium in which growth occurred at a faster rate. No phasing of DNA was reported at this time, although the data indicate that some may in fact have occurred. In this case, phasing of nuclear division was reported. Subsequent studies (Maaløe, 1961) have shown that during the transition period in which the synthetic rate is eventually increased, the cell population becomes heterogeneous, some cells commencing to synthesize DNA at a more rapid rate than others. These results indicate that following a "shift up" into a richer medium two classes of cells may be eventually distinguished. Maaløe (1961) has interpreted this to indicate that cells must complete a given physiological pattern of events (either the DNA replication cycle itself or some other parameter-controlling replication rate) before increasing the DNA synthetic rate. However, a phasing of synthesis as observed with ribonucleotides would indicate that the DNA cycle of individual cells may be composed of at least two components, which differ markedly in their response to nutritional changes in the medium. Such an explanation would be consistent with the observed effect of temperature on such cells.

The stimulation of DNA synthesis as a result of wounding (see Section V,B) is characterized by a stimulation of cells often far removed from the area of trauma. This suggests the ingestion by these cells of some compound or compounds capable of stimulating DNA synthesis. These may be cellular breakdown products. On the other hand, ingestion of foreign substances may give rise to a similar stimulation. Dutton *et al.* (1962) have studied the stimulation of DNA synthesis which occurs upon addition of a specific antigen to spleen cells taken from an immunized animal. It would appear probable that the results of further detailed study of this system could be applied with profit to the study of regenerating systems, especially with reference to the ability of humoral factors to stimulate *specific* tissues or organs.

Deoxynucleosides and tides have been observed to increase the rate of division and mitosis of cells (Karpfel *et al.*, 1959; Firshein and Braun, 1960; Butros, 1959; Gruelich *et al.*, 1961) as well as of DNA synthesis (Firshein, 1961). A study of the effect of deoxynucleosides on the timing of DNA synthesis in synchronized bacteria (Lark, 1960) demonstrated that addition of low concentrations of the four deoxyribosides (deoxycytidine, thymidine, deoxyguanosine, and deoxyadenosine) would destroy periodic DNA synthesis. This effect only occurred if these compounds were present intracellularly during a period of DNA synthesis (approximately 30% of the division cycle). Deoxyribosides are normally not found within such cells (see Section II), which have been shown to contain deoxynucloetides only. The observed effects were interpreted to indicate: (*a*) that during the period of DNA synthesis the control of the synthetic process could be affected by these substances, presumably as a result of their competition with nucleotides engaged in base-pairing with primer (this might extend the period of DNA synthesis); and (*b*) that the time of initiation of a period of DNA synthesis was related, in some manner, to the time at which a previous cycle had been completed. It was assumed that a lack of absolute synchrony would lead to a different response by individual cells to the competitive effects of deoxyribosides tending to further randomize the population and destroy synchrony.

### E. THE DISSOCIATION OF THE CELL DIVISION CYCLE FROM DNA SYNTHESIS

We have already noted the use of inhibitors to study the dependence of DNA synthesis on precursor production (Section II) or on synthesis of protein and RNA (Sections II and IV). Such experiments have also proved extremely useful in studying the dissociation of DNA synthesis from mitosis, or cell division.

Mercaptoethanol has been used to block the development of centrioles in fertilized sea urchin eggs, an event known to precede DNA synthesis. Nevertheless, DNA synthesis (if the agent is added in interphase) will proceed (Bucher and Mazia, 1960). On the other hand, if cells in metaphase are blocked by this agent, DNA synthesis is also blocked, indicating that DNA in the condensed form of the metaphase chromosome cannot duplicate. In this case, an interesting phenomenon may occur. The centers divide and form a quadripolar figure and upon the release of inhibition the egg divides into four, each of the nuclei receiving only half of its normal complement of DNA. Nevertheless, this appears to be adequate since such cells go on to further division and development. The DNA deficiency is made up before the next division since the cells go through a cycle in which the required number of centriolar units are restored (Mazia, 1960). It would be of interest to know whether the DNA so made, is made in two cycles of normal duration, including $G_1$, $S$, and $G_2$, or whether the cell immediately replenishes the necessary material in one big step.

As pointed out in this section (see Section V,B and C), under certain conditions, both temperature and nutritional changes can be used to dissociate the cell division cycle from DNA synthesis. Similarly, inhibitors of protein synthesis in bacteria may block cell division while allowing a DNA cycle to proceed (Maruyama and Lark, 1961), whereas an inhibitor of DNA synthesis (deoxyadenosine) will block DNA synthesis while permitting a further cycle of cell division (Lark, 1960).

Finally, repeated observations of synchronous cultures have shown that cycles of cell division and DNA synthesis may often be dissociated under conditions of continued cell growth and proliferation (Lark, 1962).

## VI. SUMMARY, CONCLUSION, AND OPINIONS

To date, the most complete body of experimental results concerning the regulation of DNA biosynthesis has been concerned with the role of precursors. From the data described above, there seems to be little doubt that production of DNA precursors may act as part of a mechanism for initiating DNA biosynthesis, and that repression of certain enzymes may lead to a lack of such synthesis. In correlating the production of DNA precursors with DNA biosynthesis two relationships may be encountered: The first, which may exist in germinal and embryonic tissue, could involve the induction of a metabolic pathway as a result of accumulation of large quantities of DNA precursors. In this situation, the sudden production of precursor may set in motion a train of events leading to the initiation of DNA biosynthesis. The second, which may exist in cells from adult organisms, in vegetative cells, and in bacteria, is one in which production of precursors occurs as a result of derepression and the machinery for producing DNA precursors is either synthesized or activated upon demand. The two systems differ only in that in the first, a stimulus applied at one end of the reaction chain forces biochemical events toward initiation of biosynthesis, whereas in the second, the initiation of biosynthesis forces a need for mechanisms which supply the necessary precursor materials. In both cases the mechanism for supplying the necessary precursors may be the same; however, in the first this mechanism may act as stimulus to initiate biosynthesis.

In no case, however, can such mechanisms exercise more than a crude control over the synthesis of DNA. It is impossible to explain the termination of synthesis and its reinitiation with no apparent disappearance of the enzymatic mechanisms evoked. Nor can the stoichiometry of DNA biosynthesis, resulting in the exact duplication of all of the material initially present, be accounted for. Finally, the rather specific pattern of time at which individual chromosomes are duplicated (discussed in Chapter II) cannot be controlled by such a mechanism.

An additional, more intimate and precise control appears to be operative which is capable of providing the stimulus for activating the machinery for precursor production.

The role of the primer in maintaining such a control is very attractive, but has yet to be investigated *in vivo*. Nevertheless, it would appear that it *could* play an all-important role in the control of biosynthesis; a compelling reason for assuming this, is that it is only at this level that the stoichiometry of DNA synthesis could be controlled.

Any controlling role of the DNA molecule in its own biosynthesis would appear to be complex indeed. The experiments discussed in Sections III and IV above would indicate that, among other things, the interaction of this macromolecule with others such as RNA and protein will play an important role in determining its availability to act in its own biosynthesis. Before such a mechanism can be established, experiments must be carried out on *in vivo* systems to establish the existence of a primer state and a means of quantitating its participation in biosynthesis. At present such experiments involve measurements carried out on material which has been extracted from cells in a given state of growth or division. However, the present state of microspectrophotometric techniques suggests that it may be possible to carry out physicochemical investigations of macromolecules *in situ.* If a role of primer in controlling biosynthesis can be established, the problems of the control of DNA biosynthesis will have only begun. It will then be necessary to establish the mechanisms whereby non-replicating material is converted to the primer state. The experiments discussed indicate that both protein and RNA synthesis may play an important role in this respect. The state of the template which directs RNA and protein synthesis may depend upon its interaction with these macromolecules. Moreover, it appears that DNA biosynthesis requires concomitant histone biosynthesis.

The timing of DNA biosynthesis appears to depend upon the termination of the previous period during which synthesis could take place. This suggests that some "timing reaction" is initiated upon cessation of DNA synthesis or upon cessation of a particular state ("primer") associated with DNA biosynthesis. The existence of such a reaction remains to be either proved or disproved, but the ability of chemical and physical changes in the environment to affect the DNA cycle of different cells to different extents suggests that such a mechanism exists and may be operative during $G_1$ but not during $S$, the period of DNA synthesis. A possible mechanism has been suggested in Section IV,A above.

The above conclusions concern intracellular control mechanisms operating at the molecular level. However, it is obvious that these in turn must be affected by the external environment of the cell and by its internal structure. The experiments discussed in Section V are only a small sample of the data available. However, they serve to illustrate the complexity of this problem. Within the cell, molecular components of the control mechanism must be brought together. Changes in the environment, often produced by neighboring cells, may initiate or halt the synthetic cycle. The period elapsing between mitosis and the onset of DNA synthesis appears to be extremely important in deciding the destiny of a cell. The most immediate effect of this decision may be the ability to synthesize DNA.

The significance of this aspect of the control of DNA biosynthesis may, perhaps, be viewed as follows: A series of molecular events within the cell control the functioning and duplication of DNA in a precise manner. The result of this control will determine whether the capacities of the cell are directed toward some particular purpose, other than duplication, or to replication of the cell itself. Changes produced by the interaction of the cell and its environment may be influential in shifting the balance of control to emphasize either proliferation or function. The control of DNA biosynthesis may thus present the key to the problem of cellular differentiation.

## ACKNOWLEDGMENTS

The author wishes to thank Drs. R. Okazaki and J. Paul for making available their results prior to publication. He is indebted to Drs. V. Bonifas, H. Stern, L. J. Tolmach, and P. M. B. Walker for their comments on the manuscript. Unpublished work of the author, presented above, was supported by grant E-1391 of the National Institutes of Health, Bethesda, Maryland.

## REFERENCES

Abbo, F. E., and Pardee, A. B. (1960). *Biochim. et Biophys. Acta* **39**, 478.
Alfert, M. (1955). *Symposium on Fine Structure of Cells, Leiden, 1954*, p. 157.
Alfert, M., and Das, N. K. (1961). *Proc. Am. Soc. Cell. Biol.* **1**, 6.
Aronson, A. I., and Spiegelman, S. (1961a). *Biochim. et Biophys. Acta* **70**, 70.
Aronson, A. I., and Spiegelman, S. (1961b). *Biochim. et Biophys. Acta* **70**, 85.
Astrachan, L., and Volkin, E. (1958). *Biochim. et Biophys. Acta* **32**, 449.
Baddiley, J., and Blumson, N. L. (1960). *Biochim. et Biophys. Acta* **39**, 376.
Barner, H. D., and Cohen, S. S. (1956). *J. Bacteriol.* **72**, 115.
Barner, H. D., and Cohen, S. S. (1957). *J. Bacteriol.* **74**, 350.
Barnum, C. P., Jardetzky, C. D., and Halberg, F. (1957). *Texas Repts. Biol. Med.* **15**, 135.
Bessman, M. J., Lehman, I. R., Simms, E. S., and Kornberg, A. (1958). *J. Biol. Chem.* **233**, 171.
Bianchi, P. A., Crathorn, A. R., and Shooter, K. V. (1961). *Proc. Intern. Congr. Biophys., Stockholm, 1961*, p. 230.
Billen, D. (1959a). *Nature* **184**, 174.
Billen, D. (1959b). *Biochim. et Biophys. Acta* **34**, 110.
Billen, D. (1960a). *Nature* **187**, 1044.
Billen, D. (1960b). *J. Bacteriol.* **80**, 86.
Billen, D. (1961). *Pathol. Biol.* **9**, 758.
Billen, D. (1962a). *Biochim. et Biophys. Acta*, in press.
Billen, D. (1962b). *Biochem. Biophys. Research Communs.* **7**, 173.
Bloch, D. P. (1958). *In* "Frontiers in Cytology" (S. L. Palay, ed.), p. 113. Yale Univ. Press, New Haven, Connecticut.
Bloch, D. P., and Godman, G. C. (1954). *J. Histochem. and Cytochem.* **2**, 455.
Bloch, D. P., and Godman, G. C. (1955). *J. Biophys. Biochem. Cytol.* **1**, 17.
Bojarski, B., and Hiatt, H. H. (1960). *Nature* **188**, 1112.
Bollum, F. J., and Potter, V. R. (1959). *Cancer Research* **19**, 561.
Bollum, F. J., Andregg, J. W., McElya, A. B., and Potter, V. R. (1960). *Cancer Research* **30**, 138.
Bonner, J., Huang, R. C., and Maheshwari, M. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 1548.
Brenner, S., Jacob, F., and Meselson, M. (1961). *Nature* **190**, 576.
Britten, R. J., McCarthy, B. J., and Roberts, R. B. (1962). *Biophys. J.* **2**, 83.
Brody, S., and Balis, M. E. (1959). *Cancer Research* **19**, 538.
Bucher, N. R. L., and Mazia, D. (1960). *J. Biophys. Biochem. Cytol.* **7**, 651.
Bucher, N. R. L., Scott, J. E., and Aub, J. C. (1951). *Cancer Research* **11**, 457.

Burgi, E., and Hershey, A. D. (1961). *J. Mol. Biol.* **3**, 458.

Burma, D., Kroger, H., Ochoa, S., Warner, R., and Weill, J. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 749.

Burns, V. W. (1961). *Exptl. Cell Research* **23**, 582.

Butros, J. M. (1959). *Exptl. Cell Research* **18**, 318.

Canellakis, E. S., and Montsavinos, R. (1958). *Biochim. et Biophys. Acta* **27**, 643.

Canellakis, E. S., Jaffe, J. J., Montsavinos, R., and Krakow, J. S. (1959). *J. Biol. Chem.* **234**, 2096.

Cavalieri, L. F., and Rosenberg, B. H. (1961a). *Biophys. J.* **1**, 317.

Cavalieri, L. F., and Rosenberg, B. H. (1961b). *Biophys. J.* **1**, 323.

Cavalieri, L. F., and Rosenberg, B. H. (1961c). *Biophys. J.* **1**, 337.

Cohen, S. S. (1947). *Cold Spring Harbor Symposium Quant. Biol.* **12**, 35.

Cohen, S. S., Barner, H. D., and Lichtenstein, J. (1961). *J. Biol. Chem.* **236**, 1448.

Crathorn, A. R., and Shooter, K. V. (1960). *Nature* **187**, 614.

Das, N. K., and Alfert, M. (1961a). *Proc. Natl. Acad. Sci. U.S.* **47**, 1.

Das, N. K., and Alfert, M. (1961b). *Proc. Am. Soc. Cell Biol.* **1**, 42.

Davidson, J. N., and Bishop, C. N. (1957). *Brit. J. Radiol.* **30**, 367.

De, D. N. (1961). *The Nucleus* **4**, 1.

Delamater, E. D. (1956). *Symposium Soc. Gen. Microbiol.* **6**, 215.

Doudney, C. O. (1960). *J. Bacteriol.* **79**, 122.

Doudney, C. O. (1961a). *Biochem. Biophys. Research Communs.* **4**, 218.

Doudney, C. O. (1961b). *Biochem. Biophys. Research Communs.* **5**, 405.

Doudney, C. O., and Billen, D. (1961). *Nature* **190**, 545.

Doudney, C. O., and Haas, F. L. (1961). *Genetics* **45**, 1481.

Dounce, A. L. (1959). *Ann. N.Y. Acad. Sci.* **81**, 794.

Dounce, A. L., Sarkar, N. K., and Kay, E. R. (1961). *J. Cellular Comp. Physiol.* **57**, 47.

Draculic, M., and Errera, M. (1959). *Biochim. et Biophys. Acta* **31**, 459.

Dutton, R. W., and Pearce, J. D. (1962). *Nature* **194**, 93.

Dutton, R. H., Eady, J. D., and Bulman, H. N. (1962). *16th Ann. Symposium on Fundamental Cancer Research, Houston, Texas, 1962,* in press.

Edwards, J. L., Koch, A. L., Youcis, P., Freese, H. L., Laite, M. B., and Donalson, J. T. (1960). *J. Biophys. Biochem. Cytol.* **7**, 273.

Eisenstadt, J. M., Kayeyama, T., and Novelli, G. D. (1962). *Proc. Natl. Acad. Sci. U.S.* **48**, 652.

Erikson, R. O. (1948). *Am. J. Botany* **35**, 729.

Falcone, G., and Szybalski, W. (1956). *Exptl. Cell Research* **11**, 486.

Feinendigen, L. E., Bond, V. P., and Painter, R. B. (1961). *Exptl. Cell Research* **22**, 381.

Finamore, F. J., and Volkin, E. (1958). *Exptl. Cell Research* **15**, 405.

Firket, H., and Verly, W. G. (1958). *Nature* **181**, 274.

Firshein, W. (1961). *J. Bacteriol.* **82**, 169.

Firshein, W., and Braun, W. (1960). *J. Bacteriol.* **79**, 246.

Flaks, J. G., and Cohen, S. S. (1957). *Biochim. et Biophys. Acta* **25**, 667.

Forro, F., Jr., and Wertheimer, S. A. (1960). *Biochim. et Biophys. Acta* **40**, 9.

Foster, T., and Stern, H. (1958). *Science* **128**, 653.

Foster, T., and Stern, H. (1959). *J. Biophys. Biochem. Cytol.* **5**, 187.

Fottrell, P., and Paul, J. (1961). *Biochem. J.* **80**, 17.

Fox, M. S., and Hotchkiss, R. D. (1960). *Nature* **187**, 1002.

Franzen, J. S., and Binkley, S. B. (1961). *J. Biol. Chem.* **236**, 515.

Friedkin, M. (1959). *In* "The Kinetics of Cellular Proliferation" (F. Stohlman, ed.), p. 97. Grune & Stratton, New York.

Gall, J. G. (1959). *J. Biophys. Biochem. Cytol.* **5**, 295.

Gallant, J., and Suskind, S. A. (1962). *Bacteriol. Proc.* **61**, 181.

Gallant, J., and Suskind, S. R. (1962). *Biochim. et Biophys. Acta* **55**, 627.

Geiduschek, E. P., Nakamoto, T., and Weiss, S. B. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 1405.

Gelfant, S. (1960). *Ann. N.Y. Acad. Sci.* **90**, 536.

Gelfant, S. (1961). *Proc. Am. Soc. Cell Biol.* **1**, 67.

Goldstein, A., and Brown, B. J. (1962). *Biochim. et Biophys. Acta* **53**, 19.

Goldstein, A., Goldstein, D. B., Brown, B. J., and Chou, S. (1959). *Biochim. et Biophys. Acta* **36**, 163.

Grant, P. (1958). *J. Cellular Comp. Physiol.* **52**, 227.

Gray, E. D., Weissman, S. M., Richards, J., Bell, D., Keir, H. M., Smellie, R. M. S., and Davidson, J. N. (1960). *Biochim. et Biophys. Acta* **45**, 111.

Gregg, J. R., and Løtrup, S. (1955). *Biol. Bull.* **108**, 29.

Gros, F., and Gros, F. (1958). *Exptl. Cell Research* **14**, 104.

Gros, F., Hiatt, H., Gilbert, W., Kulad, C. G., Rissebrough, R. W., and Watson, J. D. (1961). *Nature* **190**, 581.

Gruelich, R. C., Cameron, I. L., and Thrasher, J. D. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 743.

Guttes, E., and Guttes, S. (1961). *Proc. Am. Soc. Cell Biol.* **1**, 79.

Hall, B. D., and Spiegelman, S. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 137.

Hamaguchi, K., and Geiduschek, E. P. (1962). *J. Am. Chem. Soc.* **84**, 1329.

Hanawalt, P. C., Maaløe, O., Cummings, D. I., and Schaechter, M. (1961). *J. Mol. Biol.* **3**, 156.

Harding, C. V., and Srinivason, B. D. (1961). *Exptl. Cell. Research* **25**, 326.

Harold, F. M., and Ziporin, Z. Z. (1958). *Biochim. et Biophys. Acta* **28**, 492.

Harold, F. M., and Ziporin, Z. Z. (1959). *Biochim. et Biophys. Acta* **29**, 439.

Hecht, L. I., and Potter, V. R. (1956). *Cancer Research* **16**, 999.

Hecht, L. I., Potter, V. R., and Herbert, E. (1954). *Biochim. et Biophys. Acta* **15**, 134.

Hershey, A. D., Dixon, J., and Chase, M. (1953). *J. Gen. Physiol.* **36**, 777.

Hiatt, H. H., and Bojarski, T. B. (1960). *Biochem. Biophys. Research Communs.* **2**, 35.

Hoff-Jørgensen, E., and Zeuthen, E. (1952). *Nature* **169**, 245.

Holzner, J. H., Barker, T., and Popper, H. (1959). *J. Natl. Cancer Inst.* **23**, 1215.

Hornsey, S., and Howard, A. (1956). *Ann. N.Y. Acad. Sci.* **63**, 915.

Hotchkiss, R. D. (1954). *Proc. Natl. Acad. Sci. U.S.* **40**, 49.

Hotta, Y., and Stern, H. (1961a). *J. Biophys. Biochem. Cytol.* **9**, 279.

Hotta, Y., and Stern, H. (1961b). *J. Biophys. Biochem. Cytol.* **11**, 311.

Howard, A., and Dewey, D. L. (1961). *Exptl. Cell Research* **24**, 623.

Howard, A., and Pelc, S. R. (1953). *Heredity* **6** (Suppl.), 261.

Hsu, and Somers, (1962). *Abstr. 2nd Meeting Am. Soc. Cell Biol.*

Iwamura, T. (1955). *J. Biochem.* (*Tokyo*) **42**, 575.

Jacob, F., and Monod, J. (1961). *J. Mol. Biol.* **3**, 318.

Josse, J. (1962). *5th Intern. Congr. Biochem., Moscow, 1961*, in press.

Kanazir, D. (1954). *Biochim. et Biophys. Acta* **15**, 592.

Karnofsky, D. A., and Basch, R. S. (1960). *J. Biophys. Biochem. Cytol.* **7**, 61.

Karnofsky, D. A., and Bevelander, G. (1958). *Proc. Soc. Exptl. Biol. Med.* **97**, 32.

Karpfel, Z., Saska, J., and Drasil, V. (1959). *Nature* **183**, 1600.

Kas, E., and Draculic, M. (1961). *Biochem. Biophys. Research Communs.* **5**, 152.
Keir, H. M., Binnie, B., and Smellie, R. M. S. (1962). *Biochem. J.* **82**, 492.
Kellenberger, E. (1960). *Symposium Gen. Microbiol.* **10**, 39.
Kellenberger, E., Lark, K. G., and Bolle, A. (1962). *Proc. Natl. Acad. U.S.*, in press.
Kimball, R. F., and Borka, T. (1959). *Exptl. Cell Research* **17**, 173.
Kjeldgaard, N. O., Maaløe, O., and Schaechter, M. (1958). *J. Gen. Microbiol.* **19**, 607.
Klenow, H. (1961). *Biochem. Biophys. Research Communs.* **5**, 156.
Kornberg, A. (1959). *Harvey Lectures* **53**, 83.
Kornberg, A., Zimmerman, S. B., Kornberg, S. R., and Josse, J. (1959). *Proc. Natl. Acad. Sci. U.S.* **45**, 772.
Kuriki, Y., and Okazaki, R. (1959). *Embryologia* **4**, 337.
Lajtha, L. G., Oliver, R., and Ellis, F. (1954). *Brit. J. Cancer* **8**, 367.
Lark, K. G. (1960). *Biochim. et Biophys. Acta* **45**, 121.
Lark, K. G. (1961). *Biochim. et Biophys. Acta* **51**, 107.
Lark, K. G. (1962). *22nd Intern. Congr. Physiol. Sci., Leiden, 1962,* p. 766.
Lark, K. G. Unpublished data.
Lark, C., and Lark, K. G. (1962). *Biochim. et Biophys. Acta* **55**, 401.
Lark, K. G., and Maaløe, O. (1956). *Biochim. et Biophys. Acta* **21**, 448.
Lark, K. G., Cavalieri, L., and Rosenberg, B. H. (1962). Unpublished data.
Laskowski, M., Sr. (1961). *In* "The Enzymes" (P. D. Boyer, H. Lardy, and K. Myerback, eds.), Vol. V, p. 123. Academic Press, New York.
Lehman, I. R., Bessman, M. J., Simms, E. S., and Kornberg, A. (1958). *J. Biol. Chem.* **233**, 163.
LePage, G. A. (1957). *J. Biol. Chem.* **226**, 135.
Lerman, L. S., and Tolmach, L. J. (1957). *Biochim. et Biophys. Acta* **26**, 68.
Levinthal, C. (1956). *Proc. Natl. Acad. Sci. U.S.* **42**, 394.
Lieberman, I., and Ove, P. (1962). *J. Biol. Chem.* **237**, 1634.
Longer, L., and Klenow, H. (1960). *Biochim. et Biophys. Acta* **37**, 33.
Maaløe, O. (1961). *Cold Spring Harbor Symposia Quant. Biol.* **26**, 45.
Maaløe, O., and Hanawalt, P. C. (1961). *J. Mol. Biol.* **3**, 144.
McDonald, B. B. (1958). *Biol. Bull.* **118**, 71.
McDonald, B. B. (1962). *J. Cell Biol.* **13**, 193.
McFall, E., and Stent, G. (1959). *Biochim. et Biophys. Acta* **34**, 580.
McLeish, J., and Sunderland, N. (1961). *Exptl. Cell Research* **24**, 527.
Main, R. K., Cole, L. J., and Ellis, M. E. (1957). *Nature* **180**, 1287.
Marmur, J., and Grossman, L. (1961). *Proc. Nat. Acad. Sci. U.S.* **47**, 778.
Maruyama, Y. (1956). *J. Bacteriol.* **72**, 821.
Maruyama, Y., and Lark, K. G. (1959). *Exptl. Cell Research* **18**, 389.
Maruyama, Y., and Lark, K. G. (1961). *Exptl. Cell Research* **25**, 161.
Maruyama, Y., and Lark, K. G. (1962). *Exptl. Cell Research* **26**, 382.
Mazia, D. (1960). *Ann. N.Y. Acad. Sci.* **90**, 455.
Mendelsohn, M. L. (1960). *J. Natl. Cancer Inst.* **25**, 485.
Mendelsohn, M. L., Dohan, F. C., Jr., and Moore, H. A., Jr. (1960). *J. Natl. Cancer Inst.* **25**, 477.
Meselson, M., and Stahl, F. W. (1958). *Proc. Nat. Acad. Sci. U.S.* **44**, 672.
Montsavinos, R., and Canellakis, E. S. (1959). *Cancer Research* **19**, 1239.
Morris, N. R., and Fischer, G. A. (1960). *Biochim. et Biophys. Acta* **42**, 183.
Munsch-Peterson, A. (1960). *Biochem. Biophys. Research Communs.* **3**, 392.
Nakada, D. (1960). *Biochim. et Biophys. Acta* **44**, 243.
Nakada, D., and Ryan, F. J. (1961). *Nature* **189**, 398.
Nathans, D., and Lipmann, F. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 497.

Neidhardt, F. C., and Gros, F. (1957). *Biochem. et Biophys. Acta* **25**, 513.

Newton, A. A., and Wildy, P. (1959). *Exptl. Cell Research* **16**, 624.

Nomura, M., and Watson, J. D. (1959). *J. Mol. Biol.* **1**, 204.

Nygaard, O. F., Guttes, S., and Rusch, H. P. (1960). *Biochim. et Biophys. Acta* **38**, 298.

Ochoa, S., Burma, D., Kroger, H., and Weill, J. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 749.

O'Donnell, J. F., Mackal, R. P., and Evans, E. A., Jr. (1958). *J. Biol. Chem.* **233**, 1523.

Ogur, M., Minckler, S., and McClary, D. O. (1953). *J. Bacteriol.* **66**, 642.

Okazaki, R. (1959). *Biochem. Biophys. Research Communs.* **1**, 34.

Okazaki, R. (1960). *Biochim. et Biophys. Acta* **44**, 478.

Okazaki, R., and Okazaki, T. (1958). *Biochim. et Biophys. Acta* **28**, 470.

Okazaki, R., Okazaki, T., and Kuriki, Y. (1959). *Biochim. et Biophys. Acta* **33**, 289.

Okazaki, R., Okazaki, T., and Kuriki, T. (1960). *Biochim. et Biophys. Acta* **38**, 384.

Okazaki, T., and Okazaki, R. (1959). *Biochim. et Biophys. Acta* **35**, 434.

Ord, M. G., and Stocken, L. A. (1958). *Biochim. et Biophys. Acta* **29**, 201.

Overgaard-Hansen, K., and Klenow, H. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 680.

Painter, R. B., and Drew, R. M. (1959). *Lab. Invest.* **8**, 278.

Painter, R. B., Drew, R. M., and Giauque, B. G. (1960). *Exptl. Cell Research* **21**, 98.

Pardee, A. B. (1961). *Symposium Soc. Gen. Microbiol.* **11**, 19.

Pardee, A. B., and Prestidge, L. S. (1956). *J. Bacteriol.* **71**, 677.

Parizek, J., Arient, M., Dienstbier, Z., and Skoda, J. (1958). *Nature* **182**, 721.

Paul, J. (1962). Personal Communication.

Pogo, A. O., Cordero Funes, J. R., and Mordah, J. (1960). *Exptl. Cell Research* **21**, 482.

Poschkis, K. E., Bartuska, D., Zogerman, J., Goddard, J. W., and Cantarow, A. (1959). *Cancer Research* **19**, 561.

Potter, R. L., and Buettner-Janusch, T. (1958). *J. Biol. Chem.* **233**, 462.

Potter, R. L., and Schlesinger, S. (1955). *J. Am. Chem. Soc.* **77**, 6714.

Potter, R. L., Schlesinger, S., Buettner-Janusch, V., and Thompson, L. (1957). *J. Biol. Chem.* **226**, 381.

Prescott, D. M. (1960). *Exptl. Cell Research* **19**, 228.

Prescott, D. M., and Kimball, R. F. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 686.

Prescott, D. M., Bollum, F. J., and Kluss, B. C. (1962). *J. Cell Biol.* **13**, 172.

Prescott, D. M., Kimball, R. F., and Carrier, R. F. (1962). *J. Cell Biol.* **13**, 175.

Prusoff, W. H. (1959). *Biochem. Pharmacol.* **2**, 221.

Quastler, H. (1960). *Ann. N.Y. Acad. Sci.* **90**, 580.

Quastler, H. (1962). *In* "Chemical and Biological Effects of Radiation" (Haissinsky, ed.), in press.

Quastler, H., and Sherman, F. G. (1959). *Exptl. Cell Research* **17**, 420.

Reichard, P. (1960). *Biochim. et Biophys. Acta* **41**, 368.

Reichard, P., Canellakis, Z. N., and Canellakis, E. S. (1961). *J. Biol. Chem.* **236**, 2514.

Richards, B. M., Walker, P. M. B., and Deeley, E. M. (1956). *Ann. N.Y. Acad. Sci.* **63**, 831.

Roberts, R. B. (1960). *Ann. N.Y. Acad. Sci.* **88**, 752.

Roberts, R. B., Abelson, P. H., Cowie, D. B., Bolton, E. T., and Britten, R. J. (1955). Studies of biosynthesis in *E. coli. Carnegie Inst. of Wash. Publ.* **607**.

Rotherham, J., and Schneider, W. C. (1958). *J. Biol. Chem.* **232**, 853.

Rotherham, J., and Schneider, W. C. (1960). *Biochim. et Biophys. Acta* **41**, 348.

Rudner, R. (1960). *Biochem. Biophys. Research Communs.* **3**, 275.

Rueckert, R. R., and Mueller, G. C. (1960). *Cancer Research* **20**, 1584.

Sarkar, N. K. (1961). *Arch. Biochem. Biophys.* **93**, 328.

Sarkar, N. K., and Dounce, A. L. (1961). *Biochim. et Biophys. Acta* **49**, 160.

Schaechter, M. (1961). *Cold Spring Harbor Symposia Quant. Biol.* **26**, 53.

Schaechter, M., Maaløe, O., and Kjeldgaard, N. O. (1958). *J. Gen. Microbiol.* **19**, 592.

Schaechter, M., Bentzon, M. W., and Maaløe, O. (1959). *Nature* **183**, 1207.

Scherbaum, O. (1957). *Exptl. Cell Research* **13**, 24.

Scherbaum, O. H., Lauderbach, A. L., and Jalin, T. L. (1959). *Exptl. Cell Research* **18**, 150.

Schneider, J. H., Cassir, R., and Chordikion, F. (1960). *J. Biol. Chem.* **235**, 1437.

Schneider, W. C. (1955). *J. Biol. Chem.* **216**, 287.

Schneider, W. C. (1957). *J. Natl. Cancer Inst.* **18**, 569.

Schneider, W. C., and Brownell, L. W. (1957). *J. Natl. Cancer Inst.* **18**, 579.

Scott, D. B. M., and Chu, E. (1958). *Exptl. Cell Research* **14**, 166.

Seed, J. (1961). *Nature* **192**, 944.

Setlow, R. B. (1960). *Biochim. et Biophys. Acta* **39**, 180.

Setlow, R. B., and Boyce, R. (1960). *Biophys. J.* **1**, 29.

Setlow, R. B., and Setlow, S. K. (1960). *Proc. Natl. Acad. Sci. U.S.* **46**, 791.

Sherman, F. G., and Quastler, H. (1960). *Exptl. Cell Research* **19**, 343.

Sherman, F. G., Quastler, H., and Wimber, D. R. (1961). *Exptl. Cell Research* **25**, 114.

Simon, E. H. (1961). *J. Mol. Biol.* **3**, 101.

Sisken, J. E. (1959). *Exptl. Cell Research* **16**, 602.

Sisken, J. E., and Kinosita, R. (1961). *J. Biophys. Biochem. Cytol.* **9**, 509.

Smellie, R. M. S., and Eason, R. (1961). *Biochem. J.* **80**, 39P.

Smellie, R. M. S., Keir, H. M., and Davidson, J. N. (1959). *Biochim. et Biophys. Acta* **35**, 389.

Smellie, R. M. S., Gray, E. D., Keir, H. M., Richards, J., Bell, D., and Davidson, J. N. (1960). *Biochim. et Biophys. Acta* **37**, 243.

Smith, C. L., Newton, A. A., and Wildy, P. (1959). *Nature* **184**, 107.

Smith, R., and Maaløe, O. (1962). *Biochim. et Biophys. Acta*, in press.

Spiegelman, S. B., Hall, D., and Storck, R. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 1135.

Stanners, C. P., and Till, J. E. (1960). *Biochim. et Biophys. Acta* **37**, 406.

Stent, G. S., and Brenner, S. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 2005.

Stern, H. (1960a). *18th Symposium on Growth, New York, 1959,* p. 135.

Stern, H. (1960b). *Ann. N.Y. Acad. Sci.* **90**, 440.

Stern, H. (1961). *J. Biophys. Biochem. Cytol.* **9**, 271.

Stern, H. (1962). Personal communication.

Stern, H., and Hotta, Y. (1962). Personal communication.

Strauss, B., and Okubo, S. (1960). *J. Bacteriol.* **79**, 464.

Strominger, J. L., and Scott, S. S. (1959). *Biochim. et Biophys. Acta* **35**, 376.

Sueoka, N. (1960). *Proc. Natl. Acad. Sci. U.S.* **46**, 83.

Sugino, Y. (1960). *Biochim. et Biophys. Acta* **40**, 425.

Sugino, Y., Sugino, N., Okazaki, R., and Okazaki, T. (1960). *Biochim. et Biophys. Acta* **40**, 417.

Swift, H. H. (1950). *Physiol. Zoöl.* **23**, 169.

Takats, S. (1962). *Am. J. Botany* **49**, 748.

Taylor, J. H. (1960a). *Advances in Biol. and Med. Phys.* **7**, 107.

Taylor, J. H. (1960b). *Ann. N.Y. Acad. Sci.* **90**, 409.

Taylor, J. H. (1960c). *J. Biophys. Biochem. Cytol.* **7**, 455.

Taylor, J. H., and McMaster, R. D. (1954). *Chromosoma* **6**, 489.

Taylor, J. H., Woods, P. S., and Hughes, W. L. (1957). *Proc. Natl. Acad. Sci. U.S.* **43**, 122.

Taylor, J. H., Haut, W. F., and Tung, J. (1962). *Proc. Natl. Acad. Sci. U.S.* **48**, 190.

Terasima, T., and Tolmach, L. J. (1961). *Nature* **190**, 1210.

Terasima, T., and Tolmach, L. J. (1962). Unpublished results.

Utsumi, K. R. (1961). *Gann.* **52**, 203.

Vendrely, R., and Vendrely, C. (1956). *Intern. Rev. Cytol.* **5**, 171.

Walker, P. M. B., and Yates, H. B. (1952). *Proc. Roy. Soc.* **B140**, 274.

Walker, P. M. B., and Mitchison, J. M. (1957). *Exptl. Cell Research* **13**, 167.

Weissman, S. M., Smellie, R. M. S., and Paul, J. (1960). *Biochim. et Biophys. Acta* **45**, 101.

Whitfield, J. F., and Rixon, R. H. (1959). *Exptl. Cell Research* **18**, 126.

Whitmore, G. F., Stanners, C. P., Till, J. E., and Gulyas, S. (1960). *Biochim. et Biophys. Acta* **45**, 101.

Whitmore, G. F., Stanners, C. P., Till, J. E., and Gulyas, S. (1961). *Biochim. et Biophys. Acta* **47**, 66.

Williamson, D. H., and Scopes, A. W. (1960). *Exptl. Cell Research* **20**, 338.

Williamson, M. B., and Gaschlbauer, W. (1961). *J. Biol. Chem.* **236**, 1463.

Witkin, E. M. (1956). *Cold Spring Harbor Symposia Quant. Biol.* **21**, 123.

Witkin, E. M. (1959). *Proc. 10th Intern. Congr. Genet. Montreal, 1958,* p. 280.

Woodard, J., Gelber, B., and Swift, J. (1961). *Exptl. Cell Research* **23**, 258.

Young, I. E., and Fitz-James, P. C. (1959). *Nature* **183**, 372.

Zeuthen, E., and Scherbaum, O. (1954). *Proc. Symposium Colston Research Soc.* **7**, 41.

## *Chapter* V

# Molecular Mechanism of Mutations

ERNST FREESE[1]

*Department of Genetics,*
*University of Wisconsin,*
*Madison, Wisconsin*

I. Introduction . . . . . 207
  A. Classification of Heritable Variations . . . . . 207
  B. Development of Genetic Systems and Discovery of the First Mutagens 209
  C. The Genetic System of *rII* Mutants of T2 or T4 Phages . . . . 211
  D. Further Subdivision of Mutations . . . . . 212
II. Structure, Duplication, and Alteration of the Hereditary Material . . 215
  A. Structure and Replication of DNA and RNA . . . . . 215
  B. Structure and Duplication of Chromosomes . . . . . 216
  C. Classification of Base Alterations in Nucleic Acids . . . . . 216
III. Biochemistry of Mutagens and Theory of Base Alterations . . . . 217
  A. Inhibition of Nucleic Acid Precursors . . . . . 218
  B. Incorporation of Base Analogs . . . . . 219
  C. Dyes . . . . . 224
  D. Chemical Alteration of Resting Nucleic Acid . . . . . 225
  E. Radiations . . . . . 240
  F. Unclassified Mutagens . . . . . 245
  G. Spontaneous Mutations . . . . . 245
IV. The Extent of Point Mutations . . . . . 248
  A. Genetic Mapping of Mutations . . . . . 248
  B. Spontaneous and Induced Reverse Mutations . . . . . 249
  C. The Kinetics of Mutation Induction . . . . . 250
V. Mutagenic Specificity . . . . . 252
  A. The Specificity of Forward Mutations . . . . . 252
  B. The Specificity of Reverse Mutations . . . . . 253
  References . . . . . 264

## I. INTRODUCTION[2]

### A. CLASSIFICATION OF HERITABLE VARIATIONS

The concept of mutation has played an important role ever since the large variety of biological species had been attributed to evolution from a common origin. A new species was thought to come about by the gradual change of an old species. Lamarck (1809) attributed this change to both the crossing of existing forms and the inheritance of acquired characters, to which the organism had adapted during its lifetime. Darwin (1859), however, explained evolution by natural selection and reasoned that new forms arise by chance irrespective of whether they are useful, useless, or deleterious to the organism. But convincing evidence, showing that acquired (adapted) characters are not indefinitely inherited, accumulated only slowly; for example, the spontaneous nature of mutations in bacteria has been proved only recently (Luria and Delbrück, 1943). Phenotypic variations of similar organisms could then be divided into those which are heritable and those which are non-heritable and for some time this proved the only experimentally verifiable distinction. During this period, the word "mutation" was synonymous with all "heritable variations."

---

[1] Present address: Laboratory of Molecular Biology, National Institutes of Health—NINDB, Bethesda, Maryland.

[2] *Abbreviations used throughout this paper:* A, adenine; C, cytosine; HMC, 5-hydroxymethylcytosine; MeC, 5-methylcytosine; G, guanine; T, thymine; H, hypoxanthine; U, uracil; X, xanthine; G-C, guanine-cytosine pair with any of the

A further classification of heritable variations became feasible only when Mendel's laws were rediscovered at the beginning of the twentieth century and when many crosses were examined genetically and cytologically (see scheme in Fig. 1). One learned to distinguish *nuclear* from *extranuclear* (*cytoplasmic*) inheritance and their variations. The heritable nuclear variations were found to involve the alteration of the chromosomes, either their *number in a nucleus* (polyploids, aneuploids, etc.) or their *information content*. Alterations of the information content could be further divided into those that occur through a *recombination* of information already present in two pieces of genetic material and those that occur by a *mutation* of the information within one chromosome and do not require the interaction of genetic material. This present meaning of the word "mutation" is much more restricted than the original one.

The bipartition into nuclear and extranuclear variations, shown in Fig. 1, is reasonable for organisms having a defined nucleus, and it is useful, because the nucleus and the cytoplasm can often be separated by natural or artificial means. But as is common to all classifications one encounters difficulties. For certain microorganisms, e.g., bacteria, a nuclear membrane may not exist and thus a division into nuclear and extranuclear variations may be without meaning. Nevertheless, both bacteria and viruses contain most of their hereditary information in one-dimensional genetic structures (genomes) which must correspond to physical entities. The existence of such a physical entity containing

---

cytosines; dCMP, deoxycytidine monophosphate; AP, 2-aminopurine; BU, 5-bromouracil; BUdR, 5-bromodeoxyuridine; EES, ethyl ethanesulfonate; HA, hydroxylamine; NA, nitrous acid.

nucleic acids has been proved for several microorganisms. We shall name this structure a chromosome in all those cases in which it contains DNA, even though the DNA may not be visible by Feulgen staining. One can then distinguish chromosomal and extrachromosomal variations; for bacteria the latter are often called cytoplasmic variations. Since bipartition of chromosomal variations into mutation and recombination remains the same as above, the subsequent discussion about mutations refers to both macro- and microorganisms.

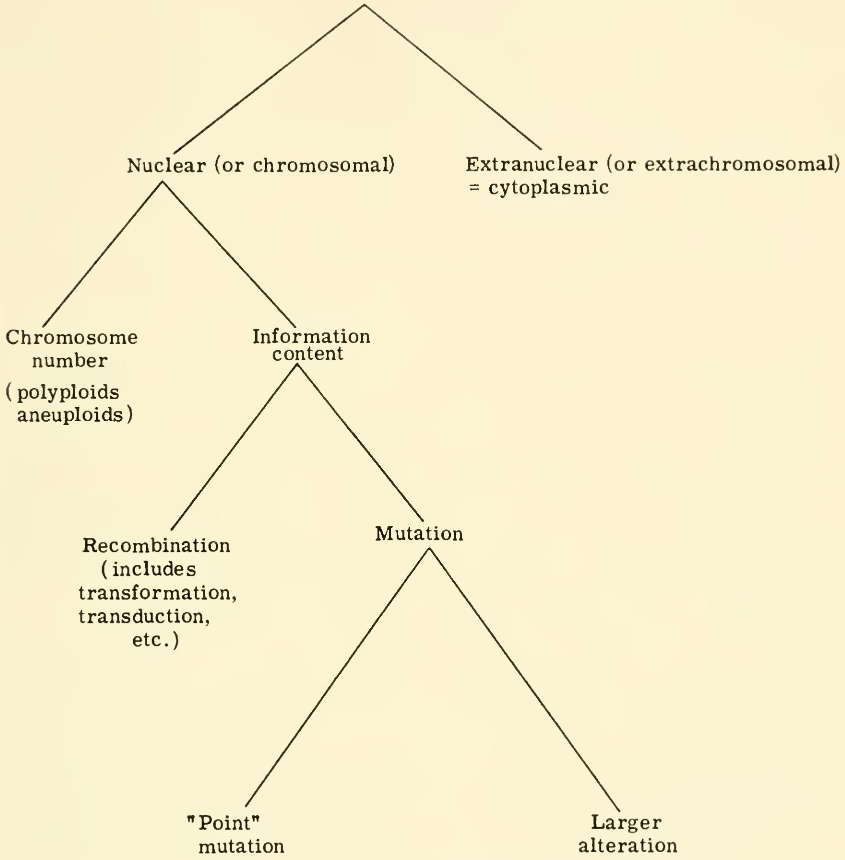

Nuclear (or chromosomal)
Extranuclear (or extrachromosomal) = cytoplasmic
Chromosome number
(polyploids aneuploids)
Information content
Recombination (includes transformation, transduction, etc.)
Mutation
"Point" mutation
Larger alteration

FIG. 1. Classification of heritable variations.

## B. DEVELOPMENT OF GENETIC SYSTEMS AND DISCOVERY OF THE FIRST MUTAGENS

The early genetic experiments were performed on higher organisms such as garden peas (Mendel, 1865) and *Oenothera* (De Vries, 1909).

De Vries discovered the first spontaneous "mutation," which was later shown to be a reciprocal translocation. These higher organisms were useful for discovering new genetic phenomena and for investigating, both cytologically and genetically, gross changes of the genetic material. With the desire to examine in detail the smaller chromosomal variations it became necessary to employ simpler genetic systems of organisms possessing a shorter generation time, thus enabling the experimenter to grow large populations and to statistically analyze genetic traits over many generations. The utilization of these systems accompanied the advancement of specific knowledge about recombination and mutation.

An important step in this direction was made by developing the genetics of the fruit fly, *Drosophila*. This genetic test system has been so refined that the frequency of x-ray-induced lethal mutations could be measured (Muller, 1927), i.e., changes in the X-chromosome that are lethal to the progeny receiving this chromosome. Radiation-induced mutations have also been observed in barley by Stadler (1928). Since then all kinds of radiations have been examined.

The first chemical mutagens were found only much later; for example, urethan + KCl which induces chromosomal breaks in *Oenothera* (Oehlkers, 1943), nitrogen and sulfur mustards (Auerbach and Robson, 1946), formaldehyde (Rapoport, 1946), and diethylsulfate (Rapoport, 1947a,b) which induce mutations in *Drosophila*. Since then many weak and a few strong chemical mutagens have been found.

The next step toward simpler genetic systems was the use of microorganisms, e.g., *Neurospora*, *Aspergillus*, yeast, and bacteria, which could be grown on a defined medium and for which biochemical mutants could be isolated, i.e., organisms that need the addition of a certain nutrient to the minimal medium. For each of these systems the induction of mutations both by radiations and chemicals has been examined. If the functional difference between the mutated and the unmutated organism is known one can distinguish *forward mutations*, which result in the loss of a biochemical function, and *reverse mutations*, which restore this function. Reverse mutations usually occur much less frequently than forward mutations but their frequency can be measured more easily, because omission from the medium of the extra nutrient, necessary for the growth of the mutant, selects for the revertant type. This selective procedure enables one to determine the frequency of revertants even in diploid cells for which the measurement of forward mutations is difficult. [For a review of recent methods and results for bacteria see Ryan (1961).]

A most important step toward an understanding of mutagenic mechanisms was the development of bacteriophage genetics. Since Hershey and Chase (1952) had shown that phage T2 injects essentially only its DNA into the bacterium it was possible to correlate genetic observations directly with alterations of the bacteriophage DNA. Other viruses have also played a significant role in this development but the most carefully analyzed tobacco mosaic virus has so far resisted all attempts to effect a genetic analysis, i.e., to obtain recombination between two mutants.

Another excellent system for mutagenic studies should be that of transforming principle (i.e., DNA). DNA could be treated by powerful reagents which cannot be used for phages or cells since they also denature protein; some of these are so specific that only one of the four bases is attacked. So far no genetic region has been examined in sufficient detail to permit a simple analysis of induced mutations both in the forward and the reverse direction.

### C. THE GENETIC SYSTEM OF *rII* MUTANTS OF T2 OR T4 PHAGES

One particular phage system, the *rII* mutants of phage T4, plays an important role in the following discussion and will therefore be described in more detail. This system has been developed by Benzer (1955) and by Chase and Doermann (1958). The phages are assayed by plating them together with about $2 \times 10^8$ sensitive bacteria, *Escherichia coli* B, (abbreviated B), on a nutrient agar plate. Each phage that infects and develops in a bacterium produces about 200 new phages which are released by lysis of the bacterium. These phages in turn infect more bacteria and thus a chain reaction produces a halo or "plaque" in the turbid bacterial meadow that is formed after several hours of incubation. Phages of "standard type" produce small plaques with fuzzy edges while *rII* mutant phages form large plaques with sharp edges. If one phage gives rise to both standard and mutant progeny a "mottled plaque" is formed and one thus has a method to detect forward mutations. Reverse mutations can be selectively measured in a different bacterium, *E. coli* K12 (λ) (abbreviated K), on which *rII* mutants cannot grow while standard type mutants or revertants can. The plaque types of the phages on the two bacterial strains are summarized in the diagram.

| | | Bacteria | |
|---|---|---|---|
| | | B | K |
| Phages | $r^+$ = Standard type | Standard | Standard |
| | *rII* type | $r$ | — |

Chase and Doermann (1958) have used this selective technique to measure the recombination frequencies between several *rII* mutants and thus to determine in two- and three-factor crosses the relative distance and order of the mutations in the *rII* region of the phage genome (chromosome). Benzer (1961) has been able to map a much larger number of mutants by using a set of *rII* mutants which contain a large alteration within the *rII* region. Such a set is shown in Fig. 2. Each *rII* mutation can thus be rapidly located in one or more of the segments demarcated by the alteration, if the mutants are crossed, by a simple spot test, against each mutant in the set. This ingenious trick reduces the amount of labor tremendously. A genetic map of spontaneous mutants is shown in Fig. 3. The same method has been used to determine rapidly the relative location of a large number of mutants induced by different means and to decide which of them could recombine with each other and which could not. Results of such examinations will be described in this paper.

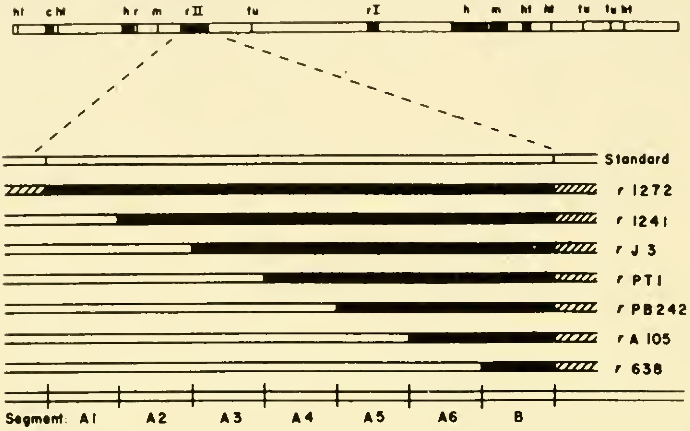


FIG. 2. At the top, the *rII* region is shown compared with the entire genetic map of the phage. This map is a composite of markers mapped in T4 and the related phage T2. Seven segments of the *rII* region are defined by a set of "deletions" beginning at different points and extending to the right band (and possibly beyond, as indicated by shading). (From Benzer, 1961, p. 404.)

### D. FURTHER SUBDIVISION OF MUTATIONS

In the attempt to subdivide mutations further, "chromosomal mutations" and "gene mutations" have been defined. But the word "gene" referred here to the older concept of a gene being at the same time the functional, recombinational, and mutational unit. Recent examinations of the genetic fine structure have proven that a functional unit contains many mutational units and can be finely subdivided by recombination;

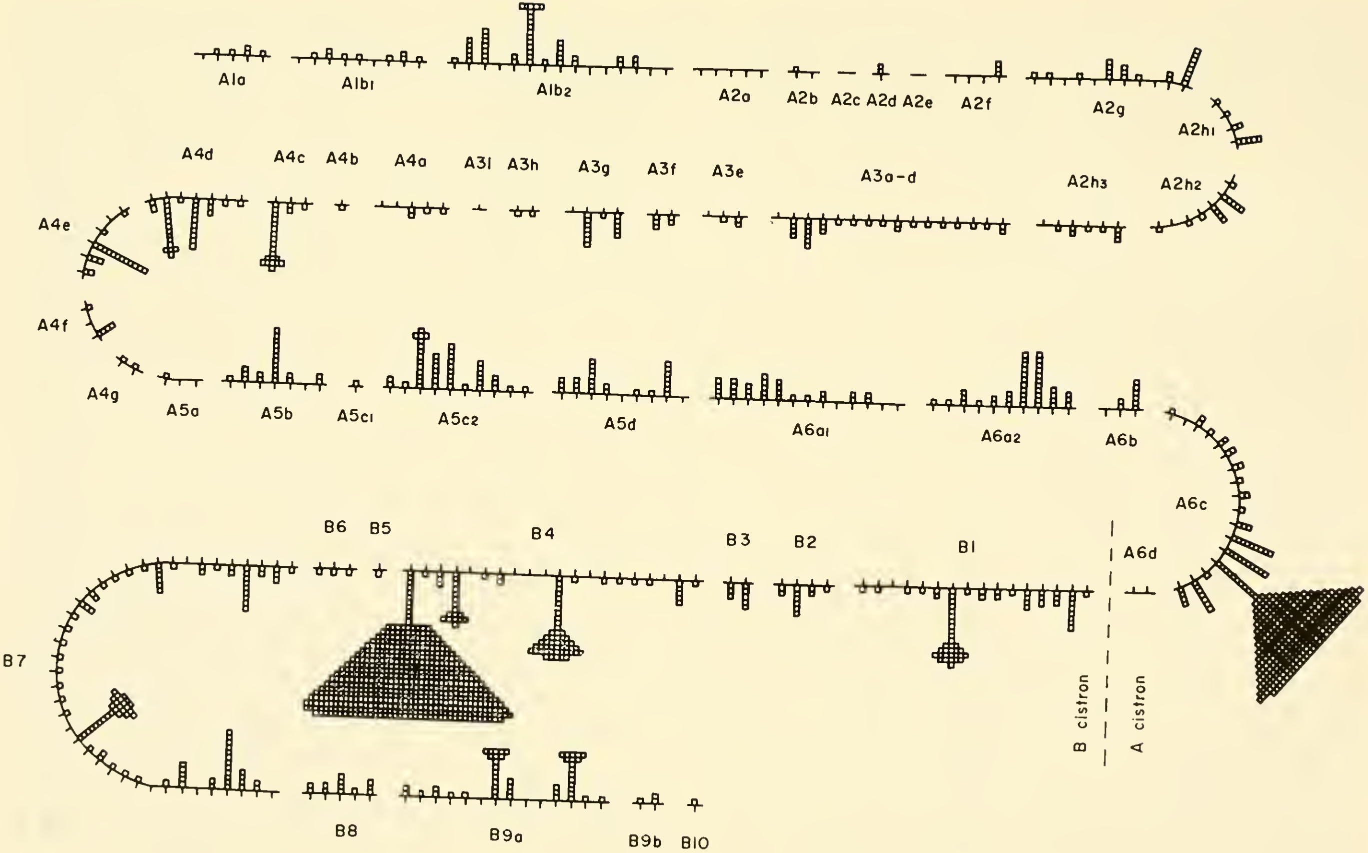

A1a
A1b1
A1b2
A2a
A2b
A2c
A2d
A2e
A2f
A2g
A2h1
A4d
A4c
A4b
A4a
A3l
A3h
A3g
A3f
A3e
A3a-d
A2h3
A2h2
A4e
A4f
A4g
A5a
A5b
A5c1
A5c2
A5d
A6a1
A6a2
A6b
A6c
B6
B5
B4
B3
B2
B1
A6d
B7
B cistron
A cistron
B8
B9a
B9b
B10

FIG. 3. Topographic map of the *rII* region for spontaneous mutations in phage T4. (From Benzer, 1961, p. 410.)

the best illustration of this fact is Fig. 3. One may thus identify the word "gene" with only one concept, and best with a functional unit (even that requires certain qualifications). "Gene mutations," then, are mutations that alter only one functional unit while "chromosomal mutations" affect several functional units at the same time. In order to define a functional unit operationally one can use the "*cis-trans*" test by which mutations that cannot cross-feed are attributed to one "cistron" (Benzer, 1957). For this test two homologous chromosomes, each carrying one mutation (*trans*-position), are placed in a common cytoplasm; when this combination does not show the *normal* functional activity the two mutations are said to belong to the same cistron. The corresponding *cis* test, in which both mutations are on one chromosome while the homologous chromosome is non-mutant, serves to establish that normal functional activity is obtained when one chromosome is not mutated. Sometimes the *cis-trans* test is not sufficient to demarcate the boundaries of a functional unit, for example, when cross-feeding is not possible because the two homologous chromosomes are too distant. Another complication is illustrated by the case in which it is not possible to decide by qualitative growth tests whether the functional activity of the heterocaryon is normal or subnormal and caused by another repair mechanism, i.e., "complementation" (see Demerec and Hartmann, 1959). In these cases it is necessary to measure quantitatively the functional (usually enzymatic) activity. Nevertheless, a suitably refined *cis-trans* test provides, in most cases, a practical definition of a functional unit (e.g., the genetic material which determines the structure of a peptide).

As another subdivision of mutations which is useful for detailed genetic experiments, one can distinguish "large alterations" and "point mutations." A mutation is a *large alteration* if the corresponding mutant cannot recombine, and produce the standard type, with at least two other mutants which in turn are able to recombine with each other. A chromosomal mutation is a large alteration; but also a gene mutation can be a large alteration, for example, if a larger piece of a gene (DNA) is missing. In contrast, one would like to define a point mutation as one in which only one nucleotide pair has been altered (including deletion or addition). Since it is impossible to prove at the present time that a given mutant has undergone such a small change the following operational definition will be used: A *point mutation* is one which does not give evidence of a larger alteration in the above-mentioned crosses with as many independently isolated mutants of the same phenotype as are available, and usually it does revert spontaneously or by induction. Obviously the precise definition of the "point" depends both on the number and kind of mutants crossed and on the smallest distance on the genetic map for which recombinants are still detectable experimentally or possible molecularly. The unit of recombination may be as small as two adjacent nucleotides, since Helinski and Yanofsky (1961) have found two very closely linked but recombining mutations of the tryptophan synthetase region for which the same amino acid is changed in the corresponding protein. The condition of reversibility is introduced in the above definition since practically all mutants of small genetic extent do revert and since one imagines that the change of a single nucleotide pair can revert occasionally. Yet, the reversion frequency of some mutations (e.g., deletions, insertions) may be unobservably small in the genetic system used. The lack of reversion does not prove that more than one nucleotide pair has been altered. In fact, some *rII* mutants were thought to be irreversible until their reversion was observed by induction. Conversely, a mutation should not be called a point mutation merely because it reverts. Unless one knows that it has been induced by a chemical which induces nothing but point mutations in other genetically well-studied systems a strong argument cannot be made, for in some cases the revertants may be the result of suppressor mutations (see Section V,B).

## II. STRUCTURE, DUPLICATION, AND ALTERATION OF THE HEREDITARY MATERIAL

We wish to determine which molecular groups of the hereditary material are changed in a mutation and how these changes come about. But before we can discuss molecular changes we need to know the molecular structure of the hereditary material itself, i.e., the molecular structure of the chromosome and the mode of its replication.

### A. STRUCTURE AND REPLICATION OF DNA AND RNA

The molecular structure of DNA, as proposed by Watson and Crick (1953), is now well established, thanks to the careful studies by Wilkins (1956), Spencer (1959), and Miles (1961). Watson and Crick (1953) also proposed that DNA duplicates by the separation of the two complementary strands and the formation of an exact new complement along each parental strand. This hypothesis is supported by studies with density labeling (Meselson and Stahl, 1958), with autoradiography (see Taylor, Chapter II), and with DNA synthesis in cell-free systems (see Bessman, Chapter I). I shall regard this mode of DNA duplication as an established theory.

Single-stranded, information-carrying DNA and RNA probably replicate also by such a complementary base-pairing mechanism. Hence most of the following considerations should also apply to mutations that occur in such systems.

### B. STRUCTURE AND DUPLICATION OF CHROMOSOMES

Since the hereditary information of chromosomes is presumably carried by DNA it would be desirable to know how the DNA is arranged in a chromosome. One possibility would be to assume that a chromosome contains just one long DNA molecule. But as the DNA apparently has to unwind during its duplication it seems more likely that occasionally one of the two DNA strands is (or becomes) interrupted in order to provide a swivel. The other DNA strand at the swivel site could be either continuous or interrupted by some non-DNA material. Details about the structure and duplication of chromosomes can be found in the article by Taylor (Chapter II). In view of the present lack of knowledge about the precise chromosomal structure it seems wiser to refrain from trying to explain large chromosomal alterations in molecular terms and to limit ourselves to the small changes that are limited to a small region of DNA.

### C. CLASSIFICATION OF BASE ALTERATIONS IN NUCLEIC ACIDS

All possible kinds of base alterations in DNA are outlined in Fig. 4. The corresponding changes of single-stranded nucleic acid can be

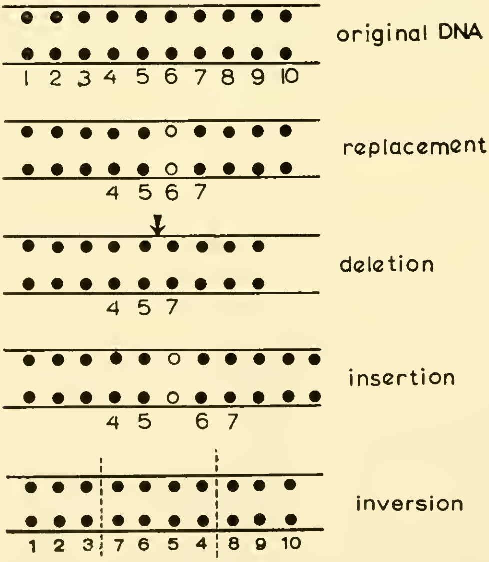


FIG. 4. Classification of the base pair changes in DNA. Each circle (●) represents a base. A hollow circle (○) is a base different from that in the original standard type DNA.

obtained by covering one of the two strands. Each of the classes of base alteration could be further subdivided, because mutagens may preferentially attack one or the other of the bases or base pairs. So far only the specificity of agents inducing replacements is sufficiently understood to warrant such a subdivision; for reasons that will become clear later it is useful to distinguish "transitions" and "transversions" (Freese, 1959b) as shown in the tabulation.

| Original base | Ultimate base | Name of base change |
|---|---|---|
| Purine | Other purine | Transition |
| Pyrimidine | Other pyrimidine | Transition |
| Purine | Any pyrimidine | Transversion |
| Pyrimidine | Any purine | Transversion |

For double-stranded DNA the possible base pair changes are the following:

$$\begin{matrix} \mathrm{A} & & \mathrm{G}\text{[3]} \\ \mathrm{T} & \longleftrightarrow & \mathrm{C} \end{matrix} = \text{transitions}$$

$$\begin{matrix} \mathrm{A} & \longleftrightarrow & \mathrm{C} \\ \mathrm{T} & & \mathrm{G} \\ \updownarrow & & \updownarrow \\ \mathrm{T} & \longleftrightarrow & \mathrm{G} \\ \mathrm{A} & & \mathrm{C} \end{matrix} = \text{transversions}$$

In order to find out how many and which nucleotide pairs of DNA have been altered in a mutation one would like ideally to make a sequence analysis of both normal and mutant DNA or at least to determine chemically that portion of DNA which has been changed. At present this is not possible; instead we have to proceed indirectly by comparing genetic and chemical observations on the effects of different mutagens on DNA.

## III. BIOCHEMISTRY OF MUTAGENS AND THEORY OF BASE ALTERATIONS

Apart from radiations there are many chemicals that show some mutagenic effect. It would be impossible to mention all these agents. Rather, I shall point out various categories and elaborate in more detail the mutagenic action of those agents whose chemical effects are better understood.

In principle one can distinguish agents that affect only replicating nucleic acid and others that exert their primary action on resting nucleic acid, although the ultimate fixation of the mutations may also require replication. Most agents have been tested only on cells or organisms and it is not clear whether they affect DNA directly or only indirectly, e.g., by alteration of a nucleic acid precursor which thus becomes mutagenic.

---

[3] See list of abbreviations in footnote 2.

### A. INHIBITION OF NUCLEIC ACID PRECURSORS

There are many agents that interfere with the normal synthesis of nucleic acid precursors (for review see Handschumacher and Welch, 1960); in addition, a number of mutants have been isolated that are blocked in one of the pathways toward nucleic acids. Some of the agents or mutants interfere with the synthesis of purines, others of pyrimidines, and again others are more specific and affect only the synthesis of one base, e.g., thymine, adenine, or guanine, sometimes only at the nucleotide level. Some of the inhibitors of the nucleic acid pathways that have shown a significant mutagenic effect are summarized in Table I. The

TABLE I

Mutagenic Agents and Mutations Which Interfere with the Normal Formation of Nucleic Acid Precursors

| Agent or mutant | Inhibits | Mutagenic effect observed in | Reference |
|---|---|---|---|
| 5-Aminouracil | Thymine | Escherichia coli<br>E. coli, biochemistry | Greer, 1958<br>Handschumacher and Welch, 1960 |
| Azaserine | Purines | E. coli<br>E. coli, biochemistry | Iyer and Szybalski, 1958, 1959<br>Handschumacher and Welch, 1960 |
| Benzimidazole | Purines | E. coli | Novick, 1956<br>Greer, 1958 |
| Caffeine | Purines | E. coli<br>E. coli, biochemistry<br>E. coli<br>Drosophila | Novick, 1956<br>Koch and Lamond, 1956<br>Glass and Novick, 1959<br>Andrew, 1959 |
| 8-Ethoxycaffeine | Thymine | Allium<br>Vicia faba | Kihlman, 1955<br>Levan, 1951 |
| Ethyl urethan | Purines | Oenothera<br>Drosophila | Oehlkers, 1949, 1953<br>Vogt, 1950 |
| 6-Mercaptopurine | Purines | E. coli | Greer, 1958 |
| 5-Nitroquinoxaline | Purines | E. coli | Greer, 1958 |
| Paraxanthine | Purines | E. coli | Novick, 1956 |
| Tetramethyl uric acid | Purines | E. coli | Novick, 1956 |
| Theobromine | Purines | E. coli | Novick, 1956 |
| Theophylline | Purines | E. coli | Novick, 1956 |
| Thymine deficiency | — | E. coli | Coughlin and Adelberg, 1956<br>Kanazir, 1958 |
| Uracil in excess | — | E. coli (chromosome breaks) | Deysson, 1952 |

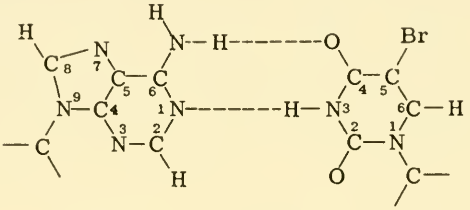
strongest mutagen in this table is azaserine which inhibits purine synthesis (see Handschumacher and Welch, 1960). But this chemical is also a very labile alkylating agent; it may therefore owe its mutagenicity to its alkylating ability rather than to its inhibitory action. Similarly urethan is both an inhibitor of pyrimidine synthesis and a mild alkylating agent. But in this case Rogers (1957) has shown that in mice thymine inhibits the production of chromosome breaks by urethan. All other growth inhibitors are rather weak mutagens in bacteria and induce chromosome breaks in higher organism. The most careful quantitative experiments about the mutation induction have been done by Novick and Szillard (see Novick, 1956). These authors have also shown that even adenine increases the frequency of mutations while the normal purine nucleosides have an anti-mutagenic effect. These observations suggest that the frequency of mutations increases because the lack of one base either causes a chromosome break, and thus a large alteration, or it increases the frequency of pairing mistakes, through which one normal nucleic acid base is incorporated in the place of another one. However, it is just as likely that the inhibitor causes the increased formation of another base analog which is the actual mutagen, perhaps by being incorporated into the nucleic acids in place of the normal bases. Many such natural base analogs have been observed in RNA [although mostly for soluble RNA, e.g., Littlefield and Dunn (1958)] and conditions of thymine starvation cause the accumulation of *N*-methyladenine in DNA (Dunn and Smith, 1955). In addition, a naturally produced mutator factor has been genetically proved to exist in a certain strain of *Salmonella typhimurium* (Miyake, 1959).

### B. INCORPORATION OF BASE ANALOGS

Apart from naturally occurring base analogs some artificial analogs are incorporated into RNA and into DNA. Here we are only concerned with those analogs that have shown a strong mutagenic effect. 5-Bromouracil (BU), 5-chlorouracil, and 5-iodouracil can replace thymine in DNA (bacteria: Weygand, Wacker, and Dellweg, 1952; Zamenhof and Griboff, 1954; phages: Dunn and Smith, 1954; Litman and Pardee, 1960; human cell lines: Szybalski and Djordjevic, 1959; Hakala, 1959; Eidinoff *et al.*, 1959), while 2-aminopurine (AP) (Wacker *et al.*, 1960a, Gottschling and Freese, 1961; Rudner, 1961) is incorporated to such a small extent that it has not been possible to determine chemically which base it replaces. 2,6-Diaminopurine is also highly mutagenic (Freese, 1959a; Benzer, 1961) but its effect has not been studied in great detail. These base analogs are much less inhibitory for bacterial and phage growth than other purine or pyrimidine analogs; yet they are much more mutagenic. For this reason their mutagenic effect has been attributed to base pairing mistakes that occur as a consequence of their incorporation into DNA (Freese, 1959a).

### 1. *5-Bromouracil (BU) or 5-Bromodeoxyuridine (BUdR)*

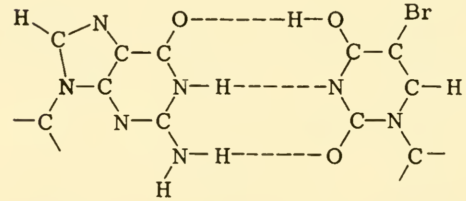
Figure 5A illustrates that BU can pair with adenine (A) as well as thymine (T) can. When T is lacking, addition of BU (or better, BUdR) to the growth medium leads to the quantitative replacement of T by BU in DNA. In most cases extensive death has been reported under such growth conditions, but when phages (Stahl, 1961) or transforming principle (Szybalski *et al.*, 1960) are kept in the absence of light very little death occurs. In other organisms BU may have side effects on vital cell constituents different from DNA. Occasionally a base pair change in DNA is induced when BU loses its hydrogen atom in the 1-position and pairs with guanine (G) instead of A (Fig. 5B). This pairing "mistake"



**Adenine**
**(normal amino state)**

**5-Bromouracil**
**(normal keto state)**

**(A)**



**Guanine**
**(normal amino state)**

**5-Bromouracil**
**(rare enol state)**

**(B)**

FIG. 5. Base pairing of 5-bromouracil (BU). A. Normal base pair of BU with A. B. Rare base pair of BU with G.

may occur either when BU enters DNA, as a "mistake in incorporation" (Fig. 6A), or after it has been incorporated into it, as a "mistake in replication" (Fig. 6B). Hence, both kinds of base pair transitions, G-C into A-T and vice versa, should be inducible with this agent. Although BU-induced mutations apparently are due to the BU incorporation, the above scheme shows that the frequency of forward mutations is not necessarily proportional to the amount of BU incorporated (Litman and Pardee, 1956) since the ratio of mistakes in incorporation to mistakes in replication may vary. For the induction of reverse mutations in bacteria such a proportionality has been shown (Strelzoff, 1962).

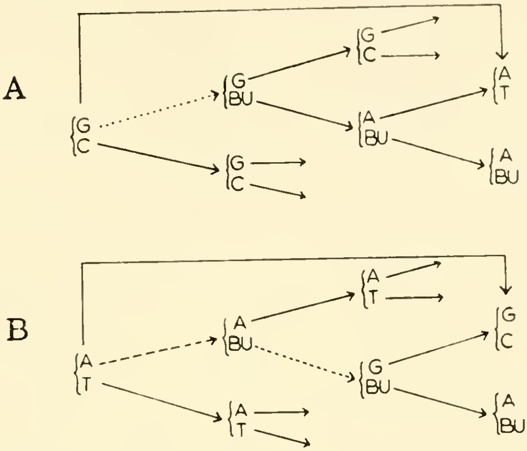


FIG. 6. The induction of base pair transitions by 5-bromouracil. **A.** Mistake in incorporation. During DNA duplication dBUTP rarely may pair "by mistake" with a G present in DNA and the dBUMP is incorporated into the new DNA strand. This initiates a base pair change since in the next DNA duplication dATP pairs with this BU, A being the "normal" complementary base of BU, and dAMP is incorporated into the new strand. Thus eventually an A-T pair appears in some of the progeny DNA in place of the original G-C pair. **B.** Mistake in replication. BU first is incorporated into DNA by pairing with its normal complementary base A. This increases the propensity of this base pair to mutate since BU undergoes pairing mistakes more frequently than T. In one of the subsequent DNA duplications dGTP may pair with BU and the incorporation of dGMP into the new DNA strand leads to the eventual appearance of a G-C pair in place of the original A-T pair.

The higher frequency of pairing mistakes which BU apparently undergoes as compared to T is probably due to the higher electronegativity of Br as compared to that of the methyl group in T. The pyrimidine ring gets poorer in electrons and apparently can lose the hydrogen in the 3-position more easily than T, thus giving rise to the ionized or tautomeric form of BU. This suggests the possibility that BU induces mistakes in incorporation (G-C → A-T) more frequently than mistakes in replication since in the latter case BU is already a member of a DNA strand and the electronegativity of its bromine may become partially reduced by electron-donating neighbor groups.

BUdR is a more efficient mutagen than BU (Freese, 1959a; Litman and Pardee, 1960) probably because it is more readily converted into the deoxynucleotide triphosphate and interferes less with the formation of uracil and cytosine. Both BU and BUdR induce point mutations (Benzer and Freese, 1958; Freese, 1959a).

The proposed mechanism of mutagenesis is supported by several other observations. When phage infected bacteria are exposed to BUdR and diluted into a medium containing excess thymidine, before any intact phages are formed, induced mutations are still observed in the lysate (Litman and Pardee, 1960). The same is observed when protein synthesis is stopped by chloramphenicol and BUdR is present only while the phage DNA multiplies (Litman and Pardee, 1959; Brenner and Smith, 1959). This shows that the induction of mutations by BUdR does not require some fixation process and does not depend on protein synthesis. The observation of mottled plaques after BUdR mutation induction shows that only one base of a DNA pair is replaced at a time and that DNA apparently duplicates by the separation of the complementary strands (Pratt and Stent, 1959), as has been assumed.

When DNA contains BU, its biological activity is much more sensitive to chemical attack. This has been observed by its sensitivity to UV (Litman and Pardee, 1960; Kozinski and Szybalski, 1959; Greer, 1960; Lorkiewicz and Szybalski, 1960), to UV and visible light (Stahl, 1961), and to X-rays (Djordjevic and Szybalski, 1960; Kaplan and Tomlin, 1960). This additional effect of BU incorporation is apparently only lethal and not mutagenic. The same applies to the increased lethality of BU-containing phages when treated by hydroxylamine (Freese *et al.*, 1961b).

### 2. *2-Aminopurine (AP)*

AP is mutagenic, apparently by a quite different mechanism. It can pair in its *normal* tautomeric form with two bases, with T by two hydrogen bonds and with C by one hydrogen bond (Freese, 1959a) (see Fig. 7). The pairing involving two hydrogen bonds should occur much more frequently than that with one hydrogen bond since in the latter case the two protonless nitrogens repel each other by their electron clouds and their relative steric position is not fixed, making it more difficult for the polymerizing enzyme to link the nucleotide to the growing DNA chain before it has separated again. Thus AP should be incorporated most frequently in the place of A which in itself would not yet be mutagenic. But whenever AP pairs with C, either in its normal tautomeric state or by two hydrogen bonds after it has undergone a tautomeric shift, a base pair change results. The frequency of the pairing "mistakes" between AP and C, per base incorporated, should be much larger than between BU and G; this may explain why, in spite of its very small incorporation, AP is highly mutagenic. Analogous to the mechanism explained for BU in Fig. 6, AP is expected to induce base pair transitions in both directions. It does induce point mutations (Freese, 1959b; Demerec, 1960).

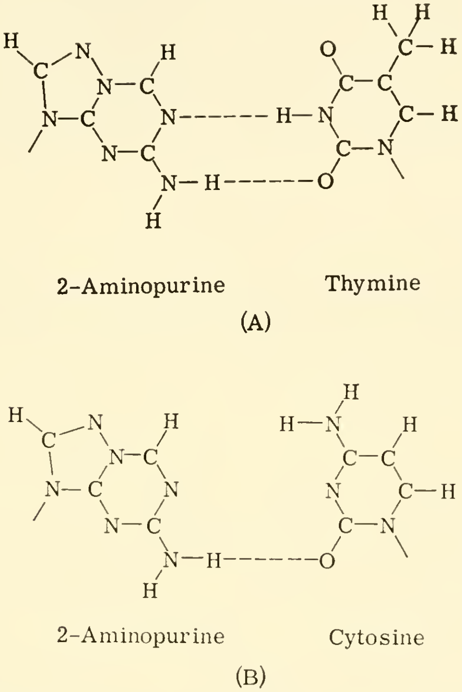
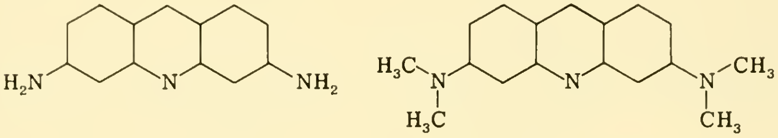


Fig. 7. Normal (A) and rare (B) base pairing of 2-aminopurine.

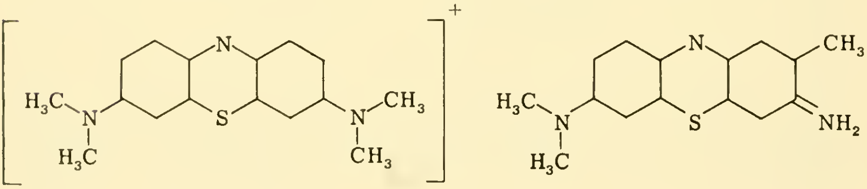
In bacteria, kinetic studies on the induction of reverse mutations by BD and AP, have been used in the attempt to distinguish the direction of the base pair transitions, by the time and pattern of mutant increase after a pulse of the mutagen (Rudner, 1960, 1961; Strelzoff, 1962). The authors observed that for some biochemical mutants the frequency of revertants increases at some time after the pulse, within one generation, and then remains constant in later generations: this indicates that the reverse mutation was induced by a mistake in incorporation. In contrast, other mutants with the same biochemical requirement showed a continued increase of revertants throughout several generations, indicating that the reverse mutation was induced by a mistake in replication. For a given mutant the time at which the first induced mutants showed up was different for AP than for BD, in the case of *Salmonella*, while it was the same in the case of *E. coli;* this may be caused by a difference in the methods used. But the over-all pattern of either cessation or continued increase of mutation induction was, for a given mutant, the same with both mutagens, in agreement with the expectations. The precise timing, however, of the appearance of revertants could also depend on the functional activity of the DNA strand in which the first pairing mistake occurs.

## C. DYES

Many different basic dyes have been used for the cytological staining of nucleic acids. Some of them, whose structures are shown in Fig. 8, have also been tested for their effect on living organisms; proflavin and acridine orange are two examples of "acridine dyes." Undoubtedly many other compounds with structures similar to those in Fig. 1 have biological effects. For example, the similarity to some aromatic carcinogens is noteworthy.



**Proflavin** **Acridine orange**



**Methylene blue** **Toluidine blue**

FIG. 8. Structure of dyes often employed in biological experiments.

A direct attachment of these dyes to nucleic acids is the probable cause of mutations. However, there are some other biological effects which will be mentioned briefly. The presence of acriflavin in the growth medium of various organisms eliminates non-chromosomal elements, e.g., the F¹ factor in *E. coli* (Hirota and Tijima, 1957; Hirota, 1960). Proflavin prevents the assembly of phage T4 (De Mars *et al.,* 1953) at concentrations at which the uninfected bacterium can still duplicate and even the phage DNA can still multiply. Certain dyes (methylene blue, toluidene blue, and proflavin) when exposed to light have a strong lethal effect on phages, provided they are inside the phage (Clifton, 1931; Yamamoto, 1956, 1958; Kaufman and Hiatt, 1959; Hiatt, 1960). This "photodynamic effect" presumably is due to the oxidation of some DNA bases, mediated by the dye and stimulated by light; for the presence of oxygen increases the effect and the presence of reducing agents (sulfhydryl groups) reduces it.

A mutagenic effect so far has been reported only for acridine dyes. Acriflavin (a mixture of proflavin and 2,8-diamino-10-methylacridinium chloride) is mutagenic for bacteria (Witkin, 1947), and this as well as other dyes are mutagenic for *Allium* (Bauch, 1948; D'Amato, 1950). Proflavin is mutagenic for phages; when it is removed at some time after phage infection normal phages are produced but an abnormally high frequency of mutants is found (De Mars, 1953; Brenner *et al.,* 1958). Many other acridine dyes similarly show a mutagenic effect on phages (Orgel and Brenner, 1961).

The molecular mechanism by which acridine dyes induce mutations is still unknown. Acridine orange and probably other acridines can stack along nucleic acids. They apparently attach to the phosphate backbone (Beers *et al.*, 1958; Bradley and Wolf, 1959) except for polyadenylic acid for which some attachment may involve the adenine bases (Steiner and Beers, 1958). Lerman (1961) has recently found that proflavine greatly increases the viscosity of DNA; from this and X-ray data he concludes that acridines can be sandwiched between two purine bases and thus cause some binding of the dye at concentrations at which the stacking is still unimportant. With these data at hand and a particular coding model in mind, Brenner *et al.* (1961) have proposed that acridines induce only deletions or insertions of one base pair in DNA.

## D. CHEMICAL ALTERATION OF RESTING NUCLEIC ACID

### 1. *Nitrous Acid*

According to Schuster (1960) nitrous acid deaminates, with decreasing frequency, the bases G, C, and A in DNA and according to Schuster and Schramm (1958) with about equal frequencies in RNA. Its mutagenic effect was first analyzed for tobacco mosaic virus and its RNA by Mundry and Gierer (1958), for bacteria (Kaudewitz, 1959), phage T2 (Vielmetter and Wieder, 1959), and many other DNA- and RNA-containing systems, including transforming principle (Litman and Ephrussi-Taylor, 1959).

Nitrous acid induces point mutations in T4 phages (Freese, 1959c; Benzer, 1961) and in yeast, i.e., *Schizosaccharomyces pombe* (Gutz, 1961).

The effect of the deaminations on the pairing properties of the different bases is depicted in Fig. 9. One can conclude (Freese, 1959c) that

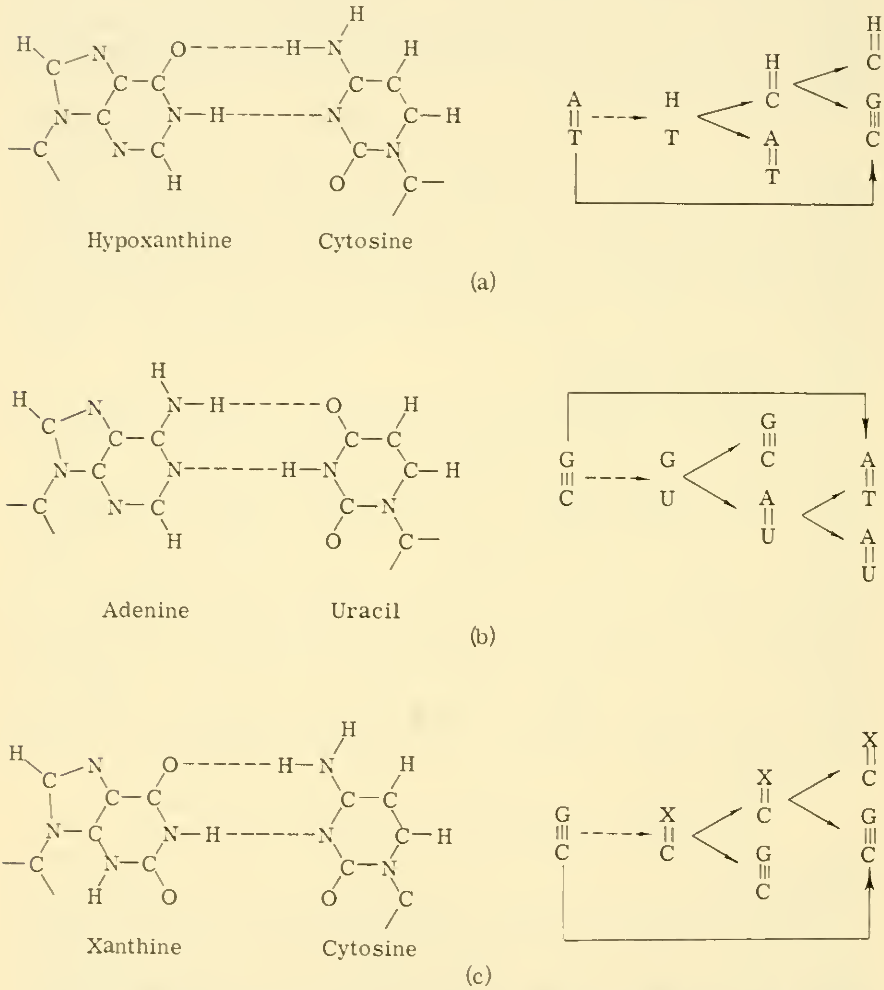


FIG. 9. Base pairing and base pair changes of deaminated bases. A is deaminated to hypoxanthine (H) which now pairs with C in place of T. C is deaminated to uracil (U) which now pairs with A in place of G. G is deaminated to xanthine (X) which *still* pairs with C, but by one less hydrogen bond.

nitrous acid should induce only, or predominantly, base "transitions" and that the deamination of G should not be mutagenic since the resulting base xanthine could still pair with cytosine. To prove this Vielmetter and Schuster (1960) compared at different pH's the rates of the deamination of the three DNA bases with the rate, per time unit, of mutation induction. They found that the deamination rates of A and C decrease with increasing pH as rapidly as the mutagenic rate while that of G decreases about one-third as fast. This shows that the deamination of G cannot be mutagenic. Since the lethal rate of nitrous acid treatment also decreases slowly with increasing pH these authors propose that the deamination of G is lethal. This conclusion is not compelling since nitrous acid exerts side reactions in addition to deamination; one can see this when dCMP is treated with nitrous acid *in vitro* and then chromatographed. Furthermore, Geiduschek (1961) has shown that nitrous acid causes cross-linking of DNA which would be almost certainly lethal for phages. The pH dependence of all these side reactions has not yet been measured. Another side reaction is the lethal effect of nitrous acid on protein; for T2 or T4 phages this is a relatively small effect, since only about 1 out of 10 lethal hits (for definition see Section IV,C) prevents the injection of phage DNA (Harm, 1960; E. B. Freese and E. Freese, 1961).

The induction of mottled plaques after nitrous acid treatment of standard type phage T2 and of non-mottled mutant plaques after treatment of phage $\phi$X174 has been used as a further proof that T2 has double-stranded and $\phi$X174 has single-stranded DNA (Tessman, 1959).

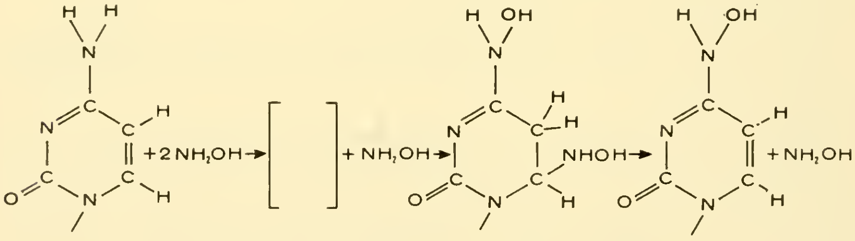
### 2. *Hydroxylamine (HA) and Hydrazine*

Both hydroxylamine (HA) and hydrazine react with certain pyrimidine bases but in spite of their chemical similarity show very different chemical specificity.

The effect of *hydrazine* on uracil has been long known (Fosse *et al.*, 1924); it has been examined again for UMP and also for CMP by Baron and Brown (1955). Hydrazine breaks the ring of uracil and cytosine, giving rise to pyrazolone and 3-aminopyrazole respectively, while the remaining urea remains attached to the sugar and in the presence of water is subsequently hydrolyzed off. Treatment of RNA by anhydrous hydrazine produces "ribo-apyrimidinic acid" free of pyrimidines (Takemura, 1957) and the treatment of DNA produces the corresponding "apyrimidinic acid" (Takemura, 1959). The author finds a small amount of thymine remaining on this DNA. In contrast, treatment of the nucleotides in a 1.5 *M* aqueous solution of pH 8.5 gives a reaction (decrease of UV absorption at 276 m$\mu$) with dTMP, UMP, and BUdR while dCMP or deoxyhydroxymethylcytidylate (dHMP) show no effect (Freese *et al.*, 1961b). It should be noted that the pK value of hydrazine is about 8.4; the pH apparently influences its chemical specificity appreciably.

The chemical effect of *hydroxylamine* (HA) on the pyrimidine bases has been observed only recently. The strongest effect (measured by the decrease of UV absorption at 260 or 276 mμ) of a 1 *M* solution at pH 7.5 is on C, CdR, dCMP, and U, UR, UMP; HMC or dHMP shows less effect, while methyl-C, T, and dTMP react too little to give any significant change of UV absorption (Freese *et al.*, 1961a,b). The difference between the methylated and unmethylated uracil and cytosine may be due to the slight electron-donating ability of the methyl group or to steric effects. 5-Bromouracil or its deoxyriboside reacts very rapidly owing to the electronegative bromine (Freese *et al.*, 1961b). Again the pH controls the chemical specificity (the pK of HA is 6.2) such that at pH 6.15 the reactivity of CMP/UMP in RNA is 30 and at pH 9.15 the ratio is 1/8 (Schuster, 1961).

The reaction of HA with U or UMP causes the irreversible loss of UV absorption in the 260–280 mμ range, breaks and removes the pyrimidine ring, producing phosphoribosyl urea and 5-isoxasolone (Schuster, 1961). The reaction with the cytosine compounds causes the loss of the 260–280 mμ absorption; but exposure to low pH return the UV absorption partially or completely. The compound obtained after exposure to low pH is not C but the hydroxylamino compound IV in Fig. 10 (Freese *et al.*, 1961b). The exact structure of the intermediate non-UV-absorbing



I II III IV

FIG. 10. Reactions of hydroxylamine with cytosine. Reaction II → III involves the liberation of an amino group. Reaction III → IV is slow at neutral but rapid at low pH.

compound obtained after reaction with DNA is not known but the analysis of 1-methylcytosine by Brown and Schell (1961) indicates that it is the dihydroxylamine compound given as formula III in Fig. 10.

Since for the DNA bases the strongest reacting base is cytosine and since the mutagenicity of this reaction can be explained, it seems probable that the major mutagenic effect of hydroxylamine in DNA is due to the alteration of C. Both the unstable compounds (II or III) or the final compound IV might be responsible for base pair changes. In order to be sure of the details the frequencies of the tautomeric states of the molecules in Fig. 10 should be known. In any event, the hydroxylamino group should be more electronegative than the amino group (because of the electronegative oxygen) and hence at least the molecule III or IV should be more frequently than C in the tautomeric state in which a hydrogen atom is on the 3-position nitrogen. This tautomeric form cannot pair with G but it can make at least one (and probably two) hydrogen bonds with A. Thus one should expect that the effect of hydroxylamine on DNA induces predominantly the base pair transitions (Freese *et al.,* 1961b).

$$\begin{matrix} G & & A \\ \cdots & \longrightarrow & \cdot\cdot \\ C & & T \end{matrix}$$

The specific effect of HA also explains why the reaction of hydroxylamine with DNA has such a small lethal effect compared to the strong mutagenicity.

The reaction of hydroxylamine with phage T4 is somewhat more complex since at low salt concentration HA mainly kills, apparently by some reaction with the phage coat, and the mutagenic effect becomes predominant only at high salt concentrations ($>1\,M$ NaCl) (Freese *et al.,* 1961a). These authors have also shown that for short reaction times most *r*-type mutants appear in form of mottled plaques while for longer treatment the fraction of pure *r* plaques increases; one would expect this since for longer treatment the likelihood increases that a second reaction occurs in the non-mutated chromosomal strand which causes it to produce non-viable or also *r*-mutant phages. HA induces point mutations in phage T4 (Freese *et al.,* 1961a).

5-Bromouracil and its deoxyriboside react very rapidly with hydroxylamine, lose their bromine and their UV absorption irreversibly. It is probable that this involves a succession of reactions but this has not yet been examined thoroughly. T4 phages containing 5-bromouracil in place of thymine are rapidly inactivated by hydroxylamine without any concomitant increase in the frequency of mutations. Similarly, hydrazine shows a much weaker mutagenic effect than hydroxylamine. It seems that the breakage of the pyrimidine ring is much more frequently lethal than mutagenic. The mutation frequencies of treatment by hydrazine are comparable to or lower than those of treatment by low pH or by ethylating agents (Freese *et al.*, 1961b). This observation again indicates that the mutagenic effect of HA is predominantly caused by its reaction with C.

### 3. *Exposure to Low pH*

When one exposes DNA to solutions below pH 4 some bases (whose pK value for the amino groups is in this pH range) become positively charged and strand separation occurs (Marmur *et al.*, 1961) which may be lethal. At or above pH 4 this effect is negligible, but low pH can affect DNA by two additional reactions which should be active even above pH 4.

*a. Depurination.* The complete removal of A and G by low pH has been reported by Tamm *et al.* (1952). Although these measurements have only been made in the pH ranges 1.6 to 4 it is clear that some depurinating effect must remain at higher pH, especially at elevated temperatures. At pH 1.6 the authors observed that G is liberated initially more readily than A. This should be measured again for pH ranges near 4.2, at which the mutagenic experiments are usually done.

No other effect of low pH on the nucleic acid bases has been reported, and extensive hydrolysis by low pH has often been used to isolate and quantitate the DNA bases.

*b. Backbone Breakage.* The sugar-phosphate backbone of DNA could be labile to treatment at low pH. This must be a minor effect since extensive treatment at pH 1.6 by the above authors gave "apurinic acid" without too much backbone breakage. However, when this apurinic acid is exposed to solutions at high pH all sugar-phosphate bonds which have no base attached to the sugar, apparently break, for the phosphate appears in the dialyzate (Tamm *et al.*, 1953). Even at neutral pH this breakage occurs, although very slowly. It is apparently caused by "$\beta$-elimination" (Bayley *et al.*, 1961). Since the biological material after exposure to low pH is returned to neutral pH some of these breaks may occur.

Exposure of T4 phages to pH 4.2 induces mutations most of which are point mutations (Freese, 1959c). It seems unlikely that such a small change can be induced by a break in the DNA backbone; a break should rather induce larger alterations or, more frequently, be lethal to the phage. Hence it is probable that the observed point mutations have been induced by depurination without any subsequent backbone breakage.

The gap left by the purine removal might be filled again before the DNA replicates. No such mechanism is known so far. Alternatively, the gap could remain empty until the DNA replicates. DNA replication could then stop at this site, which would be merely lethal, or any one of the four DNA bases could get incorporated, into the new complementary strand, or finally the gap could be left out and thus cause a small deletion (Freese, 1959c). The consequences of the incorporation of one of the four bases are illustrated in Fig. 11 for the case of a G removal; in this case one would expect both transitions and transversions.

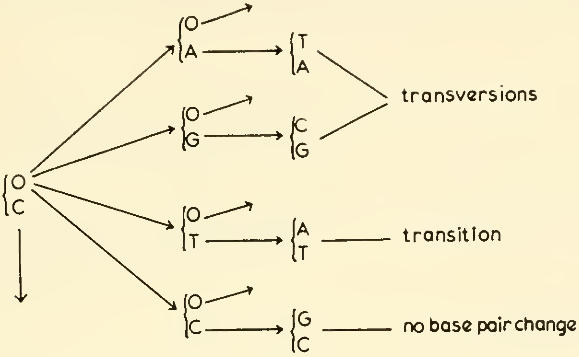


FIG. 11. Base pair replacements which could be induced by the removal of a guanine base from DNA.

### 4. *Exposure to High Temperatures*

Similar to the effect of DNA exposure to too low pH a temperature increase above the "melting temperature" causes strand separation (Doty *et al.,* 1959). Again, in addition to this effect heat apparently causes some depurination of DNA (Greer and Zamenhof, 1962) and hence must also induce chain breakage. Zamenhof and Greer (1958) showed that heating to 60°C (below the melting temperature) is mutagenic for the *E. coli* strain W6. Since the extent of these mutations has not been genetically determined it is not known whether they are "point mutations" or larger alterations. Temperature effects on the production of mutations have also been found for *Tradescantia* (Sax, 1937) and *Drosophila* (Plough, 1941).

### 5. *Alkylating Agents*

There are many agents that carry one, two, or more alkyl groups in a reactive form; these are called mono-, bi-, or polyfunctional alkylating agents. The various groups on a polyfunctional agent may either act separately or they may cause the cross-linking of molecules. Figure 12 shows the chemical structure of the most commonly used classes of alkylating agents.

All alkylating agents that prevent cell duplication have in common their reactivity with nucleophilic, i.e., slightly negatively charged, groups (Stacey *et al.*, 1958). The various chemical reactions by which alkylation can occur have been described, e.g., by Price (1958), and the relative reactivities, with respect to alkylation, of various chemical groups have been compared by Ross (1958); most reactive are sulfhydryl and thioester groups. When isolated DNA is treated in solution the reactivity of alkylating agents depends on their charge, negatively charged molecules, like chlorambucil, being least reactive. But when nucleoproteins are used (sperm heads) all alkylating agents can readily react with phosphate groups and this reactivity compares well with their lethal

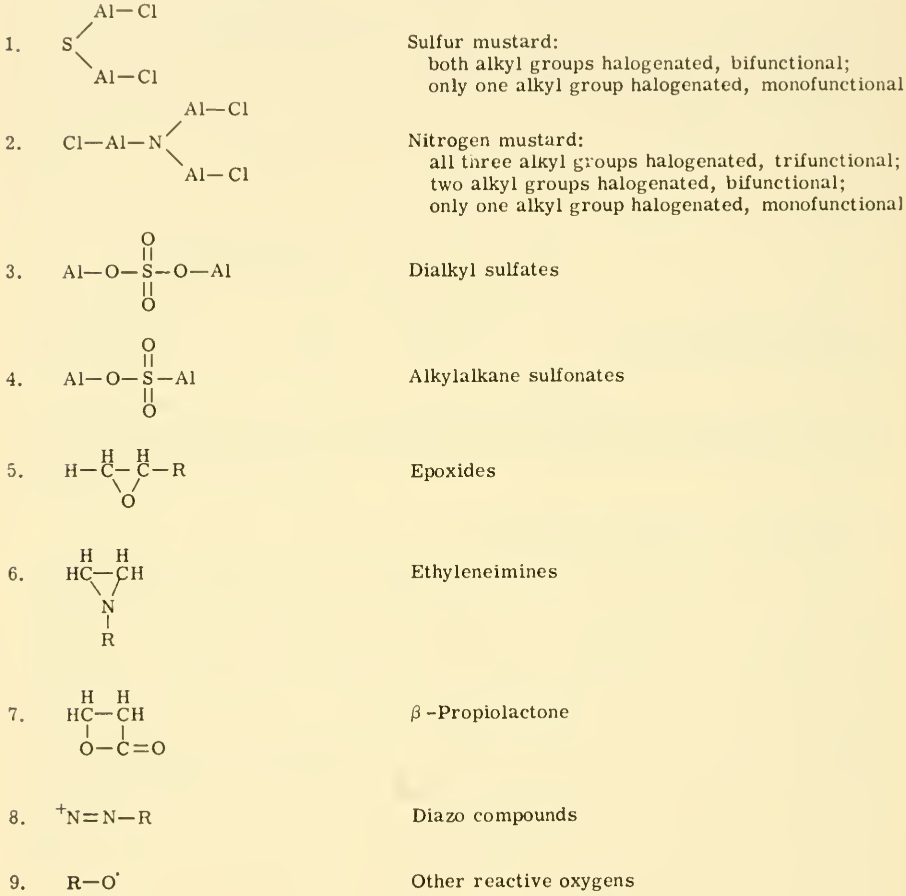

1. Sulfur mustard:
both alkyl groups halogenated, bifunctional;
only one alkyl group halogenated, monofunctional
2. Nitrogen mustard:
all three alkyl groups halogenated, trifunctional;
two alkyl groups halogenated, bifunctional;
only one alkyl group halogenated, monofunctional
3. Dialkyl sulfates
4. Alkylalkane sulfonates
5. Epoxides
6. Ethyleneimines
7. β-Propiolactone
8. Diazo compounds
9. Other reactive oxygens

Fig. 12. Different types of most commonly used alkylating agents.

effect on whole organisms (Alexander and Stacey, 1958). These authors propose that polyfunctional mustards exert their lethal effect by cross-linking of DNA molecules rather than merely by breaking the sugar-phosphate backbone of DNA. But the latter mechanism does not seem negligible since monofunctional alkylating agents also have a lethal effect.

Most alkylating agents have some mutagenic effect, whether they are mono- or polyfunctional. Their mutagenic strength and specificity, as compared to the lethal effect, depend on the structure of the alkyl group. Table II gives a list of mutagenic alkylating agents. Since they have mostly been tested on whole cells or organisms it is not known whether they act directly on DNA or the chromosomes, or whether their effect is indirect. In contrast to the necessary fixation of UV-induced mutations that requires protein synthesis no such fixation seems necessary for the mutagenic effect of at least certain alkylating agents; with diethyl sulfate and epichlorohydrin, inhibition of protein synthesis actually increases the frequency of mutations in bacteria (Strauss and Okubo, 1960).

Only few attempts have so far been made to correlate the chemical specificity of alkylating agents with their mutagenicity. An example is the trifunctional agent, triethylenemelamine, which reacts with free pyrimidine bases; the reaction product is still mutagenic with thymine but not with cytosine (Lorkiewicz and Szybalski, 1961). The authors propose that triethylenemelamine is mutagenic by altering a DNA precursor into a mutagenic base analog. This is apparently not the only effect since Fahmy and Fahmy (1958) reported that triethylenemelamine induces dominant lethals in *Drosophila* by causing large chromosomal alterations.

Most extensively analyzed with respect to both chemical effects on DNA and mutagenic effects are the methylating and ethylating agents. Such agents are dimethyl sulfate (DMS), diethyl sulfate (DES), methyl methanesulfonate (MMS), ethyl ethanesulfonate (EES), ethyl methanesulfonate (EMS). They all act as monofunctional agents even when they carry two functional groups like DMS and DES since each group alkylates separately.

DNA can be altered in five different ways by these alkylating agents (see Fig. 13).

1. The strongest reaction is the *alkylation of the phosphate groups* of nucleic acids, which has been measured for many alkylating agents (Alexander, 1952; Reiner and Zamenhof, 1957; Stacey *et al.*, 1958). The phosphate triester thus formed is unstable and for the most part hydrolyzes to return the free alkyl group (Alexander and Stacey, 1958). If

TABLE II

Alkylating Agents

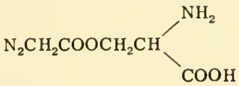
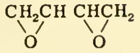
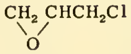
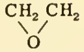
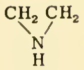
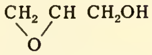
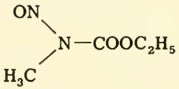
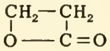
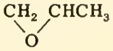
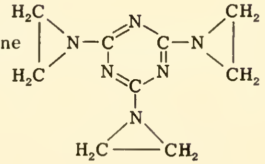
| Agent | Formula | Type | Mutagenic effect shown in | Reference |
|---|---|---|---|---|
| Azaserine |  $N_2CH_2COOCH_2CH{<}^{NH_2}_{COOH}$ | 8 | Escherichia coli | Iyer and Szybalski, 1958, 1959 |
| Bis (β chloroethyl) sulfide | $S(C_2H_4Cl)_2$ | 1 | Drosophila melanogaster | Auerbach and Robson, 1946 |
| Butyl chloroethyl sulfide | $C_4H_9SC_2H_4Cl$ | 1 | Drosophila melanogaster | Auerbach and Moser, 1950 |
| Diazomethane | $N_2CH_3$ | 8 | Drosophila melanogaster<br>Neurospora crassa | Rapoport, 1948a<br>Jensen et al., 1949 |
| Diepoxybutane |  $\underset{\backslash O /}{CH_2CH}\ \underset{\backslash O /}{CHCH_2}$ | 5 | Neurospora crassa | Kolmark and Westergaard, 1953<br>Kolmark and Giles, 1955<br>Westergaard, 1957 |
| Diethyl sulfate | $SO_2(OC_2H_5)_2$ | 3 | Drosophila melanogaster<br>Phage T4 | Rapoport, 1947<br>Loveless, 1959 |
| Dimethyl sulfate | $SO_2(OCH_3)_2$ | 3 | Drosophila melanogaster<br>Vicia faba<br>Escherichia coli | Rapoport, 1947<br>Loveless, 1951<br>Strauss, 1962 |
| Epichlorohydrin |  $\underset{\backslash O /}{CH_2\ CH}CH_2Cl$ | 5 | Neurospora crassa | Kolmark and Giles, 1955<br>Westergaard, 1957 |
| Ethylene oxide |  $\underset{\backslash O /}{CH_2\ CH_2}$ | 5 | Drosophila melanogaster | Rapoport, 1948b |
| Ethyleneimine |  $\underset{\underset{H}{\backslash N /}}{CH_2\ CH_2}$ | 6 | Drosophila melanogaster<br>Vicia faba | Rapoport, 1948b<br>Loveless, 1951 |
| Ethyl ethanesulfonate | $C_2H_5OSO_2C_2H_5$ | 4 | Phage T4 | Bautz and Freese, 1960 |

TABLE II *(continued)*

| Agent | Formula | Type | Mutagenic effect shown in | Reference |
|---|---|---|---|---|
| Ethyl methanesulfonate | $C_2H_5OSO_2CH_3$ | 4 | Drosophila melanogaster<br>Phage T4 | Fahmy and Fahmy, 1957<br>Loveless, 1959 |
| Formaldehyde | $CH_2O$ | 9 | Drosophila melanogaster<br>Drosophila melanogaster | Rapoport, 1946<br>Auerbach, 1951 |
| Glycidol |  | 5 | Neurospora crassa | Kolmark and Giles, 1955 |
| *N*-Methylbis (chloroethyl) amine | $CH_3N(C_2H_5Cl)_2$ | 2 | Drosophila melanogaster<br>Vicia faba<br>Escherichia coli | Auerbach and Robson, 1946<br>Loveless, 1951<br>Szybalski, 1958 |
| Methyl methanesulfonate | $CH_3OSO_2CH_3$ | 4 | Drosophila melanogaster<br>Phage T4 | Fahmy and Fahmy, 1955<br>Loveless, 1959 |
| *N*-Nitroso-*N*-methyl urethan |  | 8 | Vicia faba<br>Ophistoma multiannulatum | Kihlman, 1960<br>Zetterberg, 1960 |
| Phenol | $C_6H_5OH$ | 9 | Drosophila melanogaster | Hadorn and Niggli, 1946 |
| *β*-Propiolactone |  | 7 | Escherichia coli | Demerec, 1953 |
| Propylene oxide |  | 5 | Neurospora crassa | Kolmark and Giles, 1955<br>Westergaard, 1957 |
| Triethylenemelamine |  | 6 | Escherichia coli | Szybalski, 1958<br>Lorkiewicz and Szybalski, 1961<br>Fahmy and Fahmy, 1955 |

enough alkyl groups remain attached up to the time at which the DNA attempts to duplicate, the duplication might be inhibited. The attached alkyl group might conceivably (but not likely) interfere with the DNA duplication in such a way that some non-complementary base would be incorporated into the new strand.

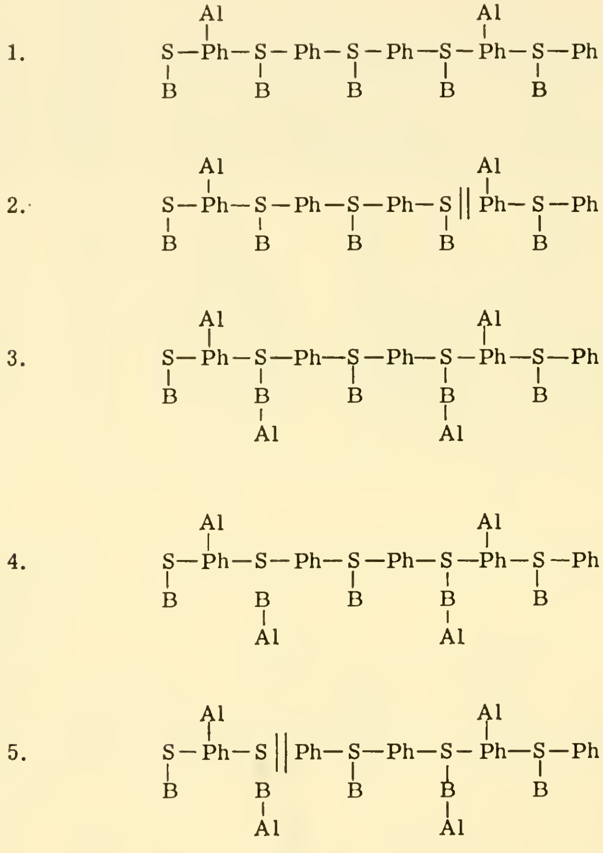


FIG. 13. Various reactions of alkylating agents with DNA (see text for explanation).

2. The *phosphate triester* can occasionally also *hydrolyze* between the sugar and the phosphate and thus the *DNA backbone* gets *broken*. The relative frequencies with which the alkyl group becomes removed or the chain is broken are not known. Chain breakage might induce larger alterations or be lethal, but probably does not induce point mutations.

3. Some of the *bases are alkylated* (see Fig. 13). Treatment of DNA

with DMS mainly produces 7-methylguanine (Reiner and Zamenhof, 1957). In addition 1-methyladenine, some 3-methyladenine and 1,3-dimethyladenine, and some cytosine derivative are formed (Brookes and Lawley, 1960a). Ethylation of the bases is less effective but some 7-ethylguanine has been found after treatment of DNA with DES (Bautz and Freese, 1960). The preferential formation of 7-alkylguanine has also been found for nitrogen mustard (Brookes and Lawley, 1960b). Lett *et al.* (1962) argue that the alkylation of the base does not occur directly but by transalkylation, the alkyl group changing from the phosphate to the base; for our purposes it is unimportant by which mechanisms the bases are alkylated. The alkylated base might inhibit DNA duplication or cause base pairing mistakes during DNA duplication.

4. *Depurination.* The alkylation of purines in the 7-position (see Fig. 14) gives rise to quaternary nitrogens which are unstable. Either the alkyl group itself hydrolyzes away from the purine or else the alkylated purine separates from the deoxyribose leaving it "depurinated." The liberation of ethylated and methylated purines from DNA

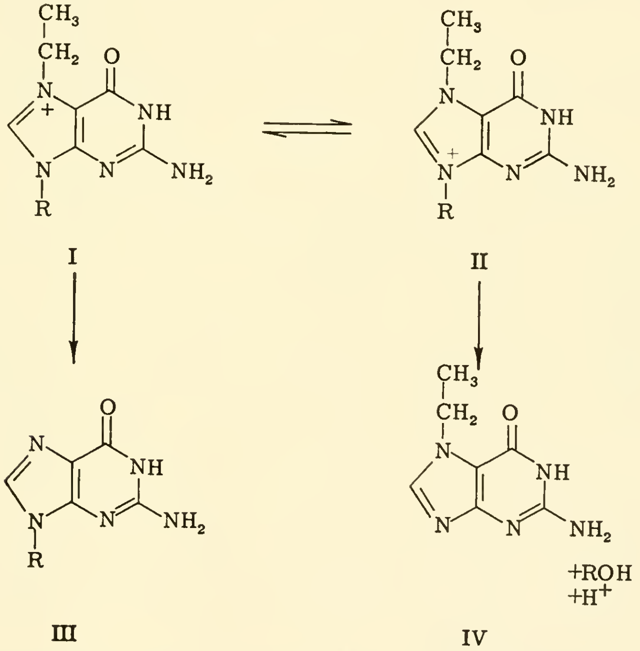


FIG. 14. Guanine removal by ethylation. I and II: Mesomeric states of 7-ethyldeoxyguanosine. III: Liberation of the ethyl-group. IV. Decay into 7-ethylguanine and deoxyribose.

has been observed (Bautz and Freese, 1960; Lett *et al.*, 1962). The gap might interfere with DNA duplication or cause the incorporation of a wrong base.

5. The depurinated DNA is labile at high pH and as shown in Section II,C,3 for the depurination by low pH, may occasionally break even at neutral pH. Again this might induce larger alterations or be lethal, but probably does not induce point mutations.

Any one of these five different effects could be lethal. Treatment of DNA with MMS or DMS rapidly methylates the phosphates and the purine bases (Reiner and Zamenhof, 1957) and leads to the depolymerization of DNA (Lett *et al.*, 1962). These agents are, however, very slightly mutagenic (per lethal hit) for bacteriophages, in contrast to the ethylating agents (Loveless, 1959; Bautz and Freese, 1960). This seems to exclude points 4 and 5 as the *major* cause of lethality, since depurination by alkylation should be as mutagenic as by low pH treatment. Of the remaining three mechanisms—1, 2, and 3—the hydrolysis of the DNA backbone (2) seems the most likely mechanism for the major cause of lethality by methylation especially since Loveless (1959) has shown that the phage titer still decreases after the treatment with MMS has been stopped by sodium thiosulfate or by dilution.

The inactivation by ethylating agents proceeds much more slowly than for methylating agents, but for phages the frequency of induced mutants per lethal hit is much larger for ethylating than for methylating agents. Since the amount of esterification of the phosphates in DNA is about the same as for methylation the high mutagenicity indicates that DNA backbone breakage by mechanism 2 is much less for ethylated than for methylated DNA. In agreement with this expectation Lett *et al.* (1962) find that EMS treatment of DNA does not influence its viscosity while MMS treatment leads to a rapid degradation. The authors interpret these results differently; they believe that mechanism 5 is mainly responsible for breakage. EMS or EES (Benzer, 1961; E. B. Freese, 1961) induce point mutations in phage T4. As for the mutations induced by low pH this finding makes it unlikely that they are caused by breaks of the DNA backbone. Even in maize ethylating agents mainly induce mutations of small genetic extent and not large alterations (Heiner *et al.*, 1960).

For the induction of mutations mechanisms 1, 3, and 4 are then to be considered. For mechanism 1 the alkyl group on the phosphate would have to interfere in some way with the pairing properties of the bases and cause the incorporation of a wrong base during DNA duplication. But the alkylated phosphate triester hydrolyzes and each removed ethyl group should decrease the frequency of mutations in treated phages. Since actually the contrary is observed (Bautz and Freese, 1960) mechanism 1 can be excluded as the major cause of mutations. Green and Krieg (1961) have furthermore shown that the labile ethyl groups cannot indirectly induce mutations inside the bacterium since a coinfected but nonalkylated phage does not reveal any mutation induction. These authors have also observed, by single burst experiments, that EMS-treated phages seem to continue throwing out normal and mutant phages even after they have replicated; this is in agreement with both mechanisms 3 and 4.

In case 3 the frequency of mutants per viable phage would be independent of subsequent incubation without DNA duplication while it might increase in case 4. Such an increase has been observed after incubation of free phages (Bautz and Freese, 1960) and for bacteria (Strauss, 1962). Besides that, mechanism 3 is expected to induce base pairing mistakes only during DNA duplication and thereby only transitions, while mechanism 4 should give essentially the same results as treatment by low pH: the removal of guanine could result in transitions as well as non-transitions. Non-transitions have actually been found (E. B. Freese, 1961) (see Table III). Hence mutations by ethylation of DNA seem to be mainly caused by the ethylation of G, which often is removed from

TABLE III

Reversion Induction of Spontaneously Reverting Mutants

| RII Mutants induced by | No. of mutants tested | Mutants found (%) Inducible to revert by AP and/or BD | Mutants found (%) Non-inducible to revert by neither AP nor BD | Approximate spontaneous background of non-reversion inducible mutants |
|---|---|---|---|---|
| 2-Aminopurine | 98 | 98 | 2 | 2 |
| 5-Bromouracil | 64 | 95 | 5 | 2 |
| Hydroxylamine | 36 | 94 | 6 | 4 |
| Nitrous acid | 47 | 87 | 13 | 15 |
| pH 5, 45°C | 115 | 77 | 23 | 15 |
| Ethyl ethanesulfonate | 47 | 70 | 30 | 10 |
| Proflavin | 55 | 2 | 98 | — |
| Spontaneous | 110 | 14 | 86 | — |

[a] From E. Freese (1961)

DNA, either during growth or before, when the treated DNA is incubated at elevated temperatures. Nevertheless, the alkylation of A seems to have some although a smaller mutagenic effect as will be seen in Section V,B,4.

It is interesting to note that the methylating agents also have some mutagenic effect: DMS is able to induce the reversion of certain *Neurospora* mutants (Westergaard, 1957) and MMS induces the reversion of certain bacterial mutants (Strauss, 1962). The frequency of these mutations could not be increased by incubation subsequent to treatment (Strauss, 1962); they may have been induced by another one of the five mechanisms mentioned or by an indirect effect.

> The emphasis on the explanation of point mutations should not distract from the fact that alkylating agents can also induce large alterations. For example, Nasrat *et al.* (1954) have shown in *Drosophila* that mustard gas induces translocations whose frequency increases with the square of the frequency of lethals; these alterations presumably are caused by a break in each of the two chromosomes involved.

Formaldehyde may also be classified as an alkylating agent. It can react with proteins (Putnam, 1953), with RNA (Fraenkel-Conrat, 1954), and with single- but not with double-stranded DNA (Staehlin, 1958). It forms monomethylol groups (—NH—$CH_2OH$) with the amino groups of the non-hydrogen-bonded nucleic acid bases; these bonds can either slowly break again or the hydroxymethyl group can react with some other group to form methylene bridges (Staehlin, 1958). Even native DNA can apparently react to some extent with formaldehyde, at least in its native state inside a phage, since Sauerbier (1960) has shown, by experiments with multiplicity reactivation, that the inactivation of phage T1 by formaldehyde is not caused by an effect on the injection mechanism. The mutagenic mechanism of formaldehyde (Rapoport, 1946; Auerbach and Moser, 1953) is not yet understood. Formaldehyde, if UV-irradiated shortly before use for mutagenic treatment, is more mutagenic (Jensen *et al.*, 1951); peroxides also increase its mutagenicity (Dickey *et al.*, 1949; Sobels, 1956). This suggests that either a radical or a peroxide is the active mutagenic agent of formaldehyde, unless the combined action of the two agents produces an additional mutagenic effect.

## E. RADIATIONS

### 1. *Ionizing Radiations*

The biological effects of X-rays and $\gamma$-rays, $\alpha$ particles, protons, high energy neutrons, and electrons have been studied so extensively that only a few review articles can here be mentioned, especially the volumes of *Radiation Biology* (1954, 1955) and Supplement 1 of *Radiation Research* (1959). The radiation can have a direct effect on the chromosome, e.g., directly break it or alter one of the DNA bases, and an indirect effect, since along the track of each particle a large number of ions or radicals remain which can initiate a chain of chemical reactions. The number of ion pairs produced in a given thickness of material, i.e., the ionization density of the radiation, depends on the nature and the kinetic energy of the particles, those of low energy giving the largest ionization density (except for neutrons which at thermal energies do not excite electrons but specifically interact with certain nuclei). The biological effect also depends on the kind of cell and on the stage of its nuclear cycle at which the irradiation is given (Marquardt, 1938). For example, chromosomes are extremely sensitive to breakage in meioitic prophase (Sax, 1941; Glass, 1955).

The inactivation of organisms often follows a first-order reaction equation

$$\frac{dN}{dt} = -\beta N \qquad \text{or} \qquad N = N_0 e^{-\beta t} = N_0 e^{-KD}$$

where $\beta$ is proportional to the intensity of the radiation. $D$ is the "dose," that is, the energy absorbed per gram of material; it is usually measured in roentgens (see Marinelli and Taylor, 1954); $K$ often is a dose independent constant describing the sensitivity of the organism. The frequency of mutants per viable organisms often increases linearly with the dose (Demerec and Sams, 1960):

$$M = \alpha D$$

Timofeeff-Ressovsky *et al.* (1935) interpreted these equations in terms of a "target theory" which states that a single "hit" of the particle on the "target," i.e., the genetic material, inactivates or mutates it. If the inactivation follows a first-order equation one calls $\beta t = KD = n$ the *number of lethal hits* because than the survival is given by the zero-order Poisson term $N/N_0 = e^{-n}$. Formally one obtains the same first-order reaction equations if the radiation product produces some chemical which can migrate from its origin to some other site in the organism or even in the reaction vessel. But the meaning of the word "target" is different and depends on the chemical or metabolic lifetime of the radiation produced chemical, i.e., whether it can get to any place or is limited to the closest surroundings of the track of the ionizing particle. If the lifetime is long the radiation produces mutagenic chemicals whose concentration depends on the number of ionizations; in that case the amount of absorbed energy, in roentgens, is the proper measure for the "dose." However, if the effective lifetime of the produced chemical (radical) is short a better measure for the dose would be the number of ionizing particles which penetrate the material, multiplied by some factor that measures the radiation efficiency of these particles. These complications arise only when one compares radiations with different ionization densities. Particles with high ionization density produce, per roentgen, more chromosomal aberrations than particles with low ionization density, probably because the chromosome has to be cut at more than one place; in contrast the induction of mutations is usually more efficient for radiations with low ionization density (Lea, 1955).

This shows that the word "single hit" can mean different reactions. It can mean a single chemical event or a single ionization, but it can also mean a single cut by a densely ionizing particle. A single-hit curve, without the precise knowledge of the chemical or physical reactions going on, does not permit one to decide whether the reaction is direct or indirect and how many chemical reactions occur until the actual inactivating or mutating event takes place.

The frequency of simple chromosome aberrations, e.g., deletions, is proportional to the dose (measured in roentgens). Hence they arise by "single hits." In contrast the more complicated ring dicentrics and exchange aberrations increase approximately with the square of the dose; they are caused by two hits (Sax, 1957). The frequency of two-hit aberrations is larger when a certain dose is given continuously than when it is fractionated, with larger rest periods. This indicates that chromosomal breaks have the chance to heal again and fewer aberrations are produced when the healing takes place before the second hit occurs (e.g., Sax, 1957). It is particularly interesting that anoxia (Wolff and Luippold, 1955) as well as chloramphenicol (Wolff, 1959), an inhibitor of protein synthesis, keep the induced breaks open longer. These observations may have a significance for the understanding of the structure of the chromosome as well as of mutation and recombination.

The importance of the indirect radiation effect was recognized even before the first chemical mutagens were known. Thus Fricke and Demerec (1937) and Friedewald and Anderson (1940) showed that rabbit papilloma virus is about 20 times more resistant in crude preparations than in a purified state. More light was shed on the indirect effect when Thoday and Read (1947) found that low oxygen concentrations reduced the frequency of chromosome breaks induced by X-radiation. This oxygen effect has been extensively studied and it is probable that radiation in the presence of oxygen causes the formation of some peroxide radical which in turn affects the frequency of chromosome breaks and maybe other mutations (see Stone, 1955). The formed peroxides may cause an increase in the frequency of chromosome breaks either because they induce breaks themselves or because they prevent the rejoining of chromosomes cut by the radiation (Baker, 1955). Reducing agents lessen the effect if added before the irradiation (Hollaender and Kimball, 1956). The major effect of the peroxides is not known; one effect on DNA could be the oxidation of the bases. In particular adenine readily reacts with peroxides giving adenine-1-*N*-oxide (Stevens *et al.*, 1958; Frederiksen and Klenow, 1960), whose nucleotides are growth inhibitors. It is not known whether the oxidation of bases in DNA is mutagenic.

In microorganisms both forward and reverse mutations have been induced by X-rays, e.g., in *Neurospora* (Giles *et al.*, 1955), in bacteria (Demerec and Sams, 1960), and in yeast (Gutz, 1961). The observations prove that X-rays also often induce point mutations.

Several phage strains are more readily inactivated in the absence than in the presence of oxygen. Alper and Ebert (1954) conclude from this and other irradiation experiments that phage is inactivated by reducing rather than by oxidizing radicals. This may be due to an effect on the phage protein rather than the DNA. (For a detailed review of the radiation effect on viruses see Gard and Maaløe, 1959; Stahl, 1959). To my knowledge there are no reports about the mutagenic effect of ionizing radiations on free viruses.

### 2. *Ultraviolet Light*

The irradiation with UV has a more specific chemical effect than ionizing radiation since the UV absorption is limited to molecules carrying conjugated double bonds and each of these molecules has a special absorption spectrum with maxima at certain wavelengths. The absorption maxima of the nucleic acid bases are in the range of 260 to 280 m$\mu$. UV of this wavelength range exerts a strong direct effect on nucleic acids. This can be seen especially well for viruses; they show no lethal effect when they are introduced into preirradiated medium but are killed by a single-hit curve when directly irradiated (McKinley *et al.*, 1926; Fisher and McKinley, 1927). The UV action spectrum of the rate of inactivation has a maximum at 260 m$\mu$ indicating that the inactivating "hits" occur directly in DNA. This is further supported by the observation that two or more UV-damaged phages can still infect the same bacterium and produce viable progeny (multiplicity reactivation, Luria, 1947). The ability of the phage to inject its DNA is inactivated much more slowly than its viability. For one needs about 10 times as many

UV hits to inactivate the *rII* function (Krieg, 1959a) or an *rII* marker (Doermann *et al.*, 1955) than to inactivate the phage. For a review see Stahl (1959).

Attempts to induce mutations by the UV-irradiation of free phages have often been unsuccessful (Weigle, 1953; Latarjet, 1954) while irradiation of the infected bacteria (Latarjet, 1949, 1954; Weigle, 1953; Weigle and Dulbecco, 1953; Tessman, 1956) or even of the bacteria alone before infection (Jacob, 1954; Fraser, 1957) readily produced UV-induced phage mutations. In some cases these "mutations" may have been caused by a radiation-induced recombination process between the phage and the bacterial genome (Stent, 1958) except for bacteriophages, like T2 or T4, which do not seem to have any similarity to the host genome (very different base ratio and HMC in their DNA). A very small direct effect has been reported by Tessman (1956) for phage T1. A significant induction of reverse mutations of a T4 *rII* mutant by UV-irradiation of free phage has been shown by Krieg (1959b). To prove the effect it was necessary to inactivate the phages quite strongly and then to infect the bacteria in high multiplicity; it is not clear whether revertants can be induced by some mistakes in multiplicity reactivation. Another direct effect is reported by Ellmauer and Kaplan (1959) for the phage K of *Serratia marcescens*.

In bacteria the possibility of an indirect induction of mutations by UV has been clearly established. Stone *et al.* (1947) showed that the UV-irradiation of bacterial media produced mutagens which induced mutations in unirradiated organisms placed in these media within less than 4 hours. The radiation product is apparently unstable and organic peroxides or radicals have been suggested as the cause of mutations. Wyss *et al.* (1948) have shown that organic peroxides are mutagenic while $H_2O_2$ in saline is not. But since UV also acts on the pyrimidine bases and the corresponding nucleotides within the bacterium, it is by no means clear how the major mutagenic effect of UV on bacteria comes about. That most UV-induced mutations somehow involve extrachromosomal material is shown by the observations (Witkin, 1959) that the frequency of UV-induced mutations is reduced when in the immediate post-irradiation period protein synthesis is blocked (by lack of amino acids or by chloramphenicol) or altered (by the incorporation of amino acid analogs) or when the formation of purines is inhibited (by non-mutagenic concentrations of caffein). In contrast the frequency of mutations is increased by the addition of nucleic acid bases. The establishment of UV-induced mutations apparently needs the presence of some unstable DNA precursors and DNA synthesis; the majority of potential UV-induced reverse mutations are established within one bacterial division following UV treatment (Witkin, 1956; Haas and Doudney, 1957).

The effect of UV has also been studied in other organisms; both mutations and chromosomal aberrations have been observed (e.g., Swanson and Stadler, 1955; Kirby-Smith and Craig, 1955; Pomper and Atwood, 1955). In yeast a mutagenic spectrum of UV-induced point mutations has been determined (Leuppold, 1962).

UV-irradiation of the pyrimidines produces unstable base analogs. The major effect is apparently the saturation of the double bond between the 5- and 6-position by water. (For a review see Shugar, 1960.) The reaction product of cytosine readily reverts at neutral pH when the irradiation is shut off. Nevertheless, this product or some derivative may be the cause of mutations. The reaction product with thymine is more stable and reverts quickly only on heating or acidification.

Recently Beukers and Berends (1960) have shown that UV-irradiation of concentrated thymine solutions as well as of DNA produces a thymine dimer. The reaction apparently goes via a higher energetic triplet state. This dimer has also been found in the DNA of UV-irradiated bacteria (Wacker *et al.*, 1960b). One can readily understand the lethal effect of the dimer formed in DNA since DNA replication is probably stopped at this place; but it is an open question whether the dimer has anything to do with UV-induced mutations.

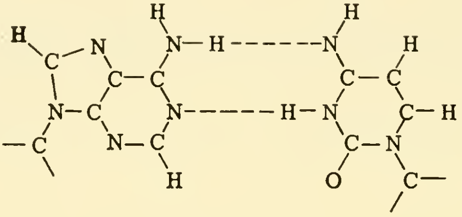
### F. UNCLASSIFIED MUTAGENS

Certain mutagens are hard to classify since so little is known about their chemical site of mutagenic attack. Such agents are, for example, manganese chloride (Demerec and Hanson, 1951), whose effect depends on the pretreatment of the bacteria, and several heavy metals (Glaess, 1956) and alkaloids (Oehlkers, 1949) in higher organisms. Another agent is 4-nitroquinoline-1-*N*-oxide, which is mutagenic for tobacco mosaic virus RNA (Endo *et al.*, 1961) and for bacteria (Mashima and Jkeda, 1958) but it is not clear whether its mutagenicity is due to the liberation of nitrous oxide or to some other effect. Also $D_2O$ was found to induce mutations slightly in bacteria (De Giovanni, 1960, 1961) and phages (Konrad, 1960) grown in its presence.

### G. SPONTANEOUS MUTATIONS

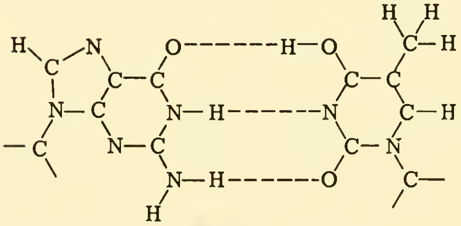
The only reasonable operational definition of "spontaneous" mutations is that these are mutations produced under "normal" growth conditions. Both the frequency and specificity of these mutations may therefore depend on the growth medium, temperature, pH, etc., which are normally employed. Experience has shown that the spontaneous frequency of mutations is approximately the same for different growth conditions usually used in a laboratory. Nevertheless, small but significant differences may be found and sometimes have been recorded as "mutagenic" or "anti-mutagenic" effect.

Some spontaneous mutations must come about by mistakes in the normal duplication of DNA or RNA. Actually it is surprising how few mistakes are made during this process. *Transitions* might come about



Adenine

Cytosine
(rare imino form)



Guanine

Thymine
(rare enol form)

FIG. 15. Rare base pairs which could occur spontaneously and would give rise to transitions. Only the cases are shown in which the pyrimidine has undergone the tautomeric shift. The same base pairs could occur when a purine has undergone a tautomeric shift or even when any one of the bases has been ionized.

by the tautomeric shift (or ionization) of one of the bases (Watson and Crick, 1953) which lead to the mistaken base pairing between A-C and, probably more frequently, G-T, as illustrated in Fig. 15. *Transversions* might come about by the mistaken pairing between two purines or two pyrimidines (Freese, 1959b). There are several such pairs possible that use one or two hydrogen bonds; in the latter case they have slightly wrong distances or angles for the two sugar (C) — base (N) bonds. Since the DNA backbone is not completely stiff it can accommodate such distortions, at least occasionally, as a mistake. The principle by which in this way transversions may come about is illustrated in Fig. 16.

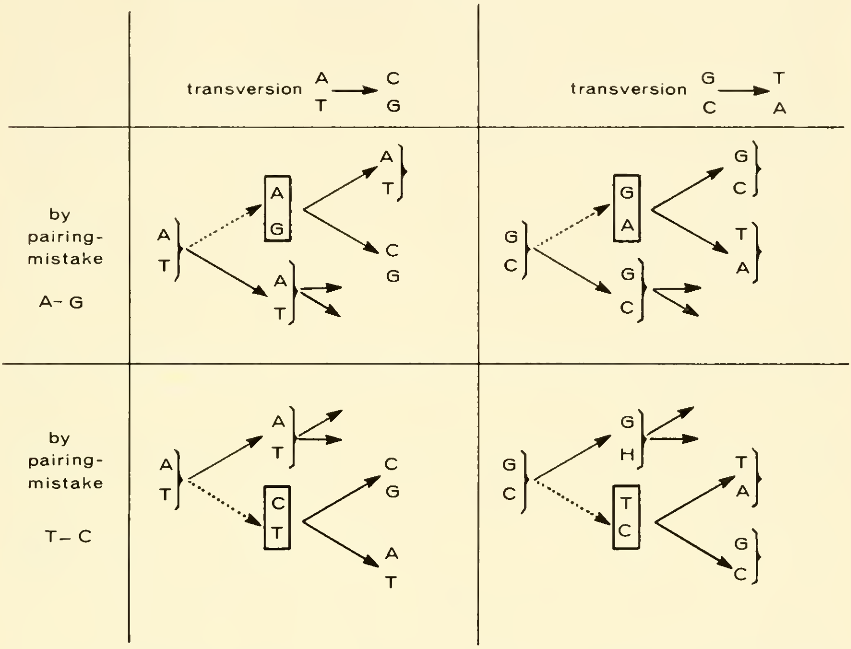


FIG. 16. Spontaneous production of transversions by rare base pairing between two purines or two pyrimidines.

Initiated by the first pairing mistake, larger alterations might be induced occasionally (Freese, 1959b). *Deletions* of one or more nucleotide pairs might also come about when one or more nucleotides in an existing DNA strand form a loop and thus do not get copied during DNA duplication, and inversely insertions might come about when by the same mechanism one or more nucleotides get copied twice (Fresco and Alberts, 1960).

Apart from these mistakes inherent in the structure and mode of replication of nucleic acids, spontaneous mutations may be caused by a variety of agents made by the cell or present in the medium. Agents of this kind are base analogs, peroxides, nitrous acid, etc., all of which are produced under some growth conditions and often are inactivated by special enzymes like catalase, peroxidase, etc. Some organisms have an especially high mutations rate (*Drosophila:* Demerec, 1937; bacteria: Treffers *et al.*, 1954); Miyake (1959) has shown that in *S. typhimurium* this property is genetically controlled and can itself mutate. Thus, "spontaneous" mutations can be controlled genetically and by the environment.

Some of these agents inside the cell act on duplicating DNA, others on resting DNA. The frequency of mutations induced by the attack of resting DNA increases with the time of exposure and may become large if resting nucleic acid is stored for a long time, as in seeds, pollen (Cartledge and Blakeslee, 1934), viruses, etc. Also in resting bacteria an increase in the frequency of mutations has been observed (Ryan, 1959). Such bacteria do not show any major turnover of DNA, but it is not clear whether or not a minor DNA turnover occurs (Ryan *et al.*, 1961).

In general it is difficult to decide which mutagenic mechanism is mainly responsible for the spontaneous mutations of a given organism. This decision is relatively easy only in special cases in which specific environmental conditions or a genetic factor greatly influence the rate. Usually one can only exclude certain mechanisms as unlikely in one of two ways. On the one hand, one can compare the genetic extent and the frequency distribution of spontaneous forward mutations with the extent and frequency distribution of mutations induced by agents of known chemical properties. This has been done for the *rII* mutants of phage T4 and very different mutability spectra have been found for spontaneous mutations and those induced by 5-bromouracil (Benzer and Freese, 1958), proflavin (Brenner *et al.*, 1958), 2-aminopurine (Freese, 1959a), and other mutagens (Benzer, 1961). One can conclude that the major spontaneous mutagenic mechanism in phage must be different from that caused by any one of the agents used for mutation induction. On the other hand, one can induce the reverse mutation of spontaneous and induced mutations by various mutagens and thus determine the properties of the base pair changes involved (see Section V).

## IV. THE EXTENT OF POINT MUTATIONS

As shown in Section III, most mutagens induce in nucleic acids the change of a single base or base pair. This does not necessarily imply that such a simple change can show up as a phenotypically observable mutation. Conceivably, several base pairs may have to be altered before such a mutation would be produced. I shall now cite several experimental results which indicate that actually a single base pair change is responsible for most mutations observed in phage T4.

### A. GENETIC MAPPING OF MUTATIONS

The classical approach to an analysis of the extent of mutations has been to cross as many independent mutants as possible with each other and to decide whether or not recombinants can arise between them. The frequency of single and double recombinants could be used to determine the sequence and relative distance of the various mutations.

As already outlined in Section I,C, the most refined genetic map is that of the *rII* region of phage T4 which Benzer (1961) has evolved using a new technique of deletion mapping. This experimental system is sensitive enough to detect any recombinant that occurs. Within the genetic region of the *rII* functional property at least 304 different mutational sites have been found and according to the Poisson distribution at least 120 more sites have not yet been detected (Benzer, 1961). Actually even more sites must exist, because so far certain mutagenic agents of a restricted specificity have been used mainly. Since the length of the *rII* region is at most $10^4$ and probably of the order of $10^3$ nucleotide pairs, each mutational site can correspond at most to a few nucleotide pairs and recombination can occur between any two such sites. Most spontaneous and practically all chemically induced mutations so far analyzed involve only one of these sites. The induced mutations include those induced by 5-bromouracil (Benzer and Freese, 1958), proflavin (Brenner *et al.*, 1958), 2-aminopurine (Freese, 1959a), nitrous acid (Freese, 1959c; Benzer, 1961), ethyl methanesulfonate (Benzer, 1961), low pH (Freese, 1959c), and hydroxylamine (Freese *et al.*, 1961a). Only about 5% of the spontaneous mutations extend over several of the presently known mutagenic sites.

### B. SPONTANEOUS AND INDUCED REVERSE MUTATIONS

If most mutations really correspond to the change of a single nucleotide pair in DNA they should occasionally mutate back to the original *genotype*, i.e., the original sequence of nucleotide pairs. Furthermore, in some cases this reversion should be inducible by some of the mutagenic agents. In contrast, a mutation which extends over several nucleotide pairs might find it difficult to revert all these nucleotide pairs simultaneously.

Actually, practically all T4-*rII* mutants induced by the aforementioned chemicals can revert to the original *phenotype* spontaneously and most of them can be induced to revert by some of the agents. Also all but 10% of the spontaneous mutants can revert spontaneously. The question is only whether these reverse mutations to the original phenotype arose by a genuine back mutation at the mutant site or rather by some other ("suppressor") mutation at a different (or even the same) genetic site.

Suppressor mutations will be treated in more detail later and it may suffice here to say that, for most spontaneous and induced mutants which have been examined in sufficient detail, revertants have been found which by functional and recombinational tests are indistinguishable from the standard type phages (Freese, 1959b; Freese *et al.*, 1961b). When a new spectrum of *rII* forward mutations was determined a few revertants behaved as standard type phages (Benzer, 1961). It is very likely, therefore, although impossible to prove unequivocally by purely genetic means, that genuine back mutations to the standard genotype exist for most *rII* mutations and hence that the original change in DNA was rather small.

Exceptions to this rule may be the mutants induced by proflavin (Crick *et al.*, 1961) which spontaneously seem to revert only, or mostly, by suppressor mutations. Otherwise the existence of suppressor mutations is unimportant for the above argument.

### C. THE KINETICS OF MUTATION INDUCTION

The ease with which most chemical reactions can be both initiated and stopped enables one to study the increase of the mutations frequency as a function of the number of nucleic bases attacked. Phage T4 is usually inactivated by a first-order reaction kinetics:

$$\frac{B}{B_0} = e^{-\beta t} = e^{-n}$$

Here $B$ is the titer of the T4 phages on bacteria $B$, $B_0$ is the titer at zero time of treatment, $\beta$ is the inactivation constant and $n$ is the number of "lethal hits," $n$ is proportional to the number of deaminated bases (as long as that number is small); it is *not identical* to the number of deaminated bases since there should be many bases which do not have any important informational task and thus can be altered without lethal effect.

$\beta$ depends on the concentration of the mutagen. If that is constant in time, as for the treatment by nitrous acid, the time of treatment itself is proportional to the number of bases attacked. When, however, the mutagen decays rather rapidly,[4] e.g., by hydrolysis as for many alkylating agents, the more adequate measure for the number of bases attacked is:

$$n = \ln \frac{B_0}{B}$$

---

[4] Once the single-hit character of mutations of a certain kind (e.g., forward mutations of a certain phenotype) is established one can use their frequency as another measure for the number of bases attacked. One then plots against this frequency the frequency of another kind of mutation. Such a plot has the advantage of being independent of the concentration of the mutagen or of an assumed constant correlation between lethal hits and bases attacked, which sometimes does not hold true (Freese *et al.*, 1961a).

If mutation induction is a one-hit phenomenon, i.e., a single chemical reaction can induce a mutation, the frequency of mutants per viable phage must increase linearly with the number of treated bases, that means with the number of lethal hits. If we call $M$ the frequency of mutants per viable phage we thus have:

$$M = M_0 + \mu n$$

where $\mu$ is a constant which depends only on the kind of mutation which one examines. It measures the probability with which the mutational site or sites are attacked by the chemical, per lethal reaction of the phage. If, in contrast, more than one chemical reaction (hit) were necessary to induce a mutation, $M$ would increase with some power of $\mu n$.

Kinetic measurements of his kind have been done with nitrous acid as mutagen for TMV (Mundry and Gierer, 1958) and phages T2 (Vielmetter and Wieder, 1959) and T4 (E. B. Freese and E. Freese, 1960), with ethyl ethanesulfonate (Bautz and Freese, 1960; E. B. Freese, 1961), with hydroxylamine (Freese *et al.*, 1961a,b), and with low pH treatment (E. B. Freese, 1961). It was always found that the frequency of *forward mutations* per visible phage (e.g., from standard type to $r$ type) increases linearly with the number of lethal hits. This indicates that each of these mutations gets induced by the chemical reaction with a single base of DNA. Not only the initial but also the ultimate change in this mutated DNA apparently involves only one nucleotide pair, since for the induction of reverse mutations again the frequency of revertants per viable phage increases linearly, as can be seen in Fig. 17 (p. 259).

The various mutagens induce about $10^{-4}$–$10^{-6}$ revertants per lethal hit for highly reversion inducible mutants. This frequency is about $10^2$–$10^3$ times smaller than that of the induction of *rII* forward mutations. Since a T2 or T4 phage has about $2 \times 10^5$ nucleotide pairs, these figures are of the right order of magnitude if one assumes that the hit of only one or any one of a few bases can induce a reverse mutation while an *rII* forward mutation can be induced by the alteration of any one of many bases within the *rII* region, and a lethal effect can be produced by the alteration of any one of even more bases. In contrast, if the alteration of more than one base were necessary to induce a reverse mutation, a two or more hit curve should be observed, with a gradually increasing slope, and the frequencies of reverse mutations per lethal hit should be much less (e.g., of the order $10^{-10}$).

It should be noted that quite a number of mutants do not respond to any *in vitro* mutagen known so far. Obviously, the above kinetic arguments do not hold for them.

## V. Mutagenic Specificity

The previous chapters have shown that each mutagen has certain chemical properties and therefore is expected to induce only certain base pair changes to the exclusion of others. This chemical specificity should correspond to a mutagenic specificity such that each mutagen can induce only certain of the possible mutations to the exclusion of others. Some selection rules should be strict and exceptions impossible, others should merely indicate a more or less pronounced mutagenic preference of a given mutagen.

A mutagenic specificity has been actually observed for the induction of reverse mutations by various mutagens. For example, Demerec (1953) reported two mutants in *E. coli* of which one (leucine-dependent) responds little to UV and much to $MnCl_2$ while the other (phenylalanine-dependent) responds in the opposite way. In *Neurospora crassa* several adenine-dependent mutants, controlling the same functional property, have been found to differ in their spontaneous and X-ray-induced revertibility (Giles, 1955; De Serres, 1958) and can be distinguished by their pattern of response to different chemical mutagens (Westergaard, 1957). A very detailed analysis of mutagenic specificity has been undertaken for the *rII* region of bacteriophage T4 which will be reported in more detail below. Several results with bacterial systems (Demerec, 1960; Rudner, 1960, 1961; Kirchner, 1960; Margolin and Mukai, 1961; Strelzoff, 1962) and with yeast (Leuppold, 1962; Gutz, 1961) corroborate and extend these findings.

### A. THE SPECIFICITY OF FORWARD MUTATIONS

Each functional property is determined by a genetic region, containing many sites whose change results in the loss of the function. It is not surprising, therefore, that forward mutations for different functional properties can be induced with similar frequency (same order of magnitude), because individual site specificities average out. But if one examines the frequency of forward mutations at one particular site quite a different picture emerges. For such an analysis a large number of independently arising mutants of the same functional property must be isolated and mapped. The results for the *rII* mutants of phage T4 show that mutations frequently recur at certain sites (hot spots). The location of these sites is specific and different for different mutagens. The highest specificity has been observed with spontaneous mutations (Benzer, 1957, 1961), next with the base analogs 5-bromouracil (Benzer and Freese, 1958) and 2-aminopurine (Freese, 1959a), and with some *in vitro* chemicals (Freese, 1959c; Benzer, 1961). Proflavin-induced mutational sites (Brenner *et al.*, 1958) are more scattered and are nearly all different from the spontaneous ones.

These observations show that the mutagenic specificity is more complicated than one might have expected from the chemical observations. Clearly the mutability of a given nucleotide pair is influenced by other structural factors. Part of this effect may be due to the different kinds of cytosine bases that are observed in DNA. 5-Methylcytosine is found in many organisms in addition to cytosine and preferentially next to G (Sinsheimer, 1955); in T2, T4, T6 phages differently glucosylated 5-hydromethylcytosine bases are observed and the extent of glucosylation depends on the neighbors (Lehman and Pratt, 1960; Kornberg *et al.*, 1961). Hydroxylamine distinguishes chemically between some of these C bases (Freese *et al.*, 1961b), hence other mutagenic mechanisms, including the spontaneous one, might do the same.

In addition, the nearest neighbors of a given base undoubtedly can have a direct influence upon its mutability, as is indicated by the sites of very high mutability mentioned above. Such effects may be produced for example by the difference in the hydrogen bond strength of the next neighbors, that of A-T pairs being less than that of G-C pairs, as it determines the ease with which the base pairs can open up, or by sterical effects of the next neighbor, purines being bulkier than pyrimidines, or by an electronic influence of the $\pi$-electron clouds of one base upon the next one.

It is conceivable that certain DNA sites have evolved by such means a special protection against spontaneous mutations and against certain mutagens.

## B. THE SPECIFICITY OF REVERSE MUTATIONS

The aim of studying the induction of reverse mutations is to divide mutations into classes which do or do not respond to a given mutagen or which respond much and little. When forward mutations, induced by a chemical of known specificity, are used the consistency of the theory of mutations can thus be checked. In addition one can hope to determine for each or at least many mutations by which base pair change they came about; this would also enable one to infer the mutagenic mechanism for agents whose chemical specificity is not sufficiently understood, as well as for the spontaneous mutations.

### 1. *Complications of the Analysis*

The interpretation of experimental results concerned with the induction of reverse mutations is complicated by two factors: the higher mutagenic specificity and suppressor mutations.

*a. The Higher Mutagenic Specificity.* The reactivity of a given nucleotide pair toward a mutagen depends on its location in the DNA, as we have already seen for the induction of forward mutations (Section V,A). Mutants containing the same type of base pair change at different sites may thus give appreciably different rates of reversion. However, this variable mutagenic specificity can only influence the height of the mutagenic response and cannot make an A-T base pair respond to a mutagen that attacks only G or C, etc. If one finds two distinct classes of mutants, those that are much and others that are little (or not) inducible by a given mutagen, the corresponding mutations must be fundamentally different. This bipartition is most simply explained by assuming that the mutations of the two classes differ by the type of base pair change by which they arose; the higher mutagenic specificity would then only control the relative strength of the response within each class.

*b. Suppressor Mutations.* Since reverse mutations occur so rarely they can be observed only on selective media on which the original mutant cannot multiply. All strains which can grow under these selective conditions are called *revertants* of the original mutant. They may arise by "back mutations" to the standard genotype, i.e., the original sequence of nucleotide pairs in DNA, or by "suppressor mutations." The latter do not return the genome to the standard state but nevertheless restore the growth of the organism partially and sometimes even completely. Such suppressor mutations may occur in the same functional region as the original mutation (*intragenic suppressor mutations*) or in a different functional region (*extragenic suppressor mutations*).

The extragenic suppressor mutations usually restore the growth of the organism only partially and then can be distinguished from back mutations on growth plates. Their presence is furthermore revealed readily by crosses of a revertant to the standard type organism. These extragenic suppressor mutations may be non-specific or specific for the original mutational site; they have interesting properties which go beyond the scope of this paper. For the *rII* region with which we are here mainly concerned extragenic suppressor mutations do not seem to exist. This follows from the fact that 10% of the spontaneous *rII* mutants do not revert at all, which excludes non-specific suppressors, and that specific extragenic suppressor mutations have so far not been found.

I shall therefore limit the further discussion to back mutations and intragenic suppressor mutations. Let us consider the case in which the functional region contains the information for a polypeptide that is the whole or a part of an enzyme. Some mutations may alter the enzyme such that the change of no amino acids in the polypeptide can repair the damage, except the restoration of the original amino acid sequence.

Such a situation may hold, e.g., for a functionally important amino acid in the "active center" of the enzyme or the presence or absence of a proline which produces a kink in the protein. In these cases only a back mutation would return full functional activity. On the other hand, many mutations may cause a change of the enzyme which can be counteracted also by a second mutation at some other genetic site. Such a suppressor mutation may return the functional property of the enzyme partially or even completely. If a given mutagen should induce this second base pair change more frequently than the back mutation, most induced revertants would contain both the original mutation and an intragenic suppressor mutation.

The question now is, how can one distinguish intragenic suppressor mutations from back mutations? If the functional property is restored only to a limited extent this distinction is easy because one observes a different phenotype. For example, many *rII* mutants plated on bacteria K produce both standard type and tiny plaques; at least the latter must be due to suppressor mutations. Proflavin mutants represent an extreme case since most revertants show plaques different from the standard phenotype. Practically all of these revertants are apparently caused by suppressor mutations and back mutations are very rare. Brenner *et al.* (1961) have proposed that proflavin induces only deletions or insertions of one base pair which rarely back mutate. Crick *et al.* (1961) have employed these mutations in order to provide some evidence for their coding theory.

The revertants can be examined further for their functional properties by plating them on different bacterial hosts of type K and B; thus one can distinguish partial revertants from revertants with standard phenotype. Occasionally one observes revertant plaques so similar to the standard type that a clear distinction is difficult; in most cases such phages have not been used for further analysis.

The next step is to determine whether any of the remaining revertants of standard phenotype still contain both the original *rII* mutation and a suppressor mutation. To this end one backcrosses revertant to standard type phages and measures the frequency of *rII* mutants before and after the cross. Many such crosses have been made, leaving out all unclear cases of intermediate plaque types, and no evidence for further suppressor mutations was found in these cases (Freese, 1959b; Freese *et al.*, 1961b; E. B. Freese, 1962). Thus it would seem that most revertants of genuine standard phenotype arose by back mutations. However, the significance of the backcrosses is limited by the fact that each standard type and revertant stock already contains spontaneous *r* mutations; the crosses thus permit one to state only that the suppressor mutation cannot be more distant from the original mutation than about 0.2 recombination units. This is only a small portion of the *rII* region, which extends over about 8 recombination units. One could reduce this possible distance even further, by examining the frequency of the original *rII* mutant, among all *r* mutants, before and after the cross (by spot tests against the deletion set); this has not been reported so far.

In another important analysis Benzer (1961) has isolated from a hot spot mutant a few revertants with standard phenotype and determined a new forward mutational spectrum for each of them. If a revertant arose by a suppressor mutation it should give a different mutational spectrum, either leaving out the original hot spot or else, if the suppressor mutation itself is highly revertible, giving only the same hot spot and no other ones. In contrast about the same mutational spectrum was found as originally present.

Even this elaborate analysis does not unequivocally prove the existence of back mutations since it is possible, though unlikely, that hot spots correspond to a group of nucleotide pairs. As a further proof it seems desirable that a correlation be made between revertants and the amino acid sequence in the corresponding protein. Interestingly enough even that would not enable a complete proof of genuine back mutation since it is possible that the same amino acid can be coded by two different nucleotide combinations. Hence the ultimate proof can only come from a partial sequence analysis of the DNA around the mutational site.

In spite of the above complications several observations on the induction of reverse mutations are clear cut and demand an explanation. When a mutagen efficiently increases the frequency of revertants for certain mutants while it leaves other mutants untouched, it seems probable that the non-inducible mutants contain a base pair change which cannot be reversed by the particular mutagen. The inducible mutants can be further examined in order to decide which of the induced revertants do or do not seem to be caused by a suppressor mutation. If there are induced revertants, which by functional and recombination tests give no indication of a suppressor mutation, it seems reasonable to assume they arose by back mutations. Hence one can conclude that the particular base pair change of these mutants can be induced to back mutate by the mutagen. It does not matter, for this argument, whether none, few, or even many suppressor mutations are induced at the same time.

### 2. *Transitions versus Other Changes*

The specificity of the induction of reverse mutations shows up most effectively for the difference between transition and non-transition mutants. For this investigation one can use mutagens, which are expected to induce only base pair transitions, but those in both directions (see Section III); such agents are AP, BUdR, and NA. All transition mutants should be inducible to revert while non-transition mutants should not respond. The two base analogs are especially useful for these studies because simple spot tests can be employed to screen a large number of mutants; such spot tests have been described for bacteria (Szybalski, 1958) and phages (Freese, 1959b). The reversion inducibility by the base analogs has been determined for many *rII* mutants and the results are given in Table III (p. 239). As expected, most of the mutants induced by the transition-inducing agents AP, BUdR, HA, and NA can be induced to revert by the base analogs. Several of them have also been tested for reversion inducibility by nitrous acid and the result was positive (E. B. Freese and E. Freese, 1961). The non-inducible mutants correspond to the background of spontaneous revertants in the phage stocks.

In contrast most of the spontaneous and proflavin-induced mutants do not respond in this way. It seems clear that most of these non-base analog inducible mutants must have been caused by base pair changes different from transitions. Also about one-third of the mutants induced by low pH or EES are of this type. It is significant that for these mutants neither back mutations nor suppressor mutations are base analog inducible (except for a few mutants whose spontaneous reversion rates are so high that induction cannot be observed).

Two-thirds of the mutants, induced by low pH and EES, can be induced to revert by base analogs; they may contain a transition, or possibly some other base pair change if the base analogs should only induce suppressor mutations for some of these mutants. But EES and low pH certainly can induce transitions; this is also shown by their ability to induce the reversion of transition mutants (see Section V,4).

### 3. *Further Subdivision of Non-Transition Mutants*

The mutants of Table III that cannot be induced to revert by the two base analogs nevertheless revert spontaneously and genetic mapping shows that they are point mutations. The question is which of these mutants contain a transversion, a deletion, or an insertion?

To answer this question, at least in part, one can measure the reversion induction by the two mutagens, low pH and EES. About 70% of the *rII* mutants, induced by these agents, can be induced to revert by the base analogs; therefore many of them seem to contain a transition. We have proposed before (Section III,D) that the gap produced by the depurination of DNA could result in both transitions and transversions, or in deletions. Since low pH and EES can induce transitions they

TABLE IV

Rates $\alpha/\beta$ of Reversion Induction for Non-Transition Mutants

| Mutant | With pH 4.2 | With EES | With nitrous acid | Spontaneous backgound of revertants $(10^{-6})$ |
|---|---|---|---|---|
| *r* 111 | 66 | 70 | ≲3 | 160 |
| *r* 207 | 57 | 55 | ≲2.5 | 69 |
| *r* 132 | 55 | 50 | ≲2 | 197 |
| *r* 157 | 40 | 50 | ≲5 | 175 |
| *r* 315 | 5.0 | 4.5 | ≲0.2 | 5.0 |
| *r* 102 | 2.2 | 3.8 | ≲0.1 | 15 |
| *r* 105 | 2.4 | 3.5 | ≲0.1 | 6.1 |
| *r* 179 | 2.4 | 3.4 | ≲0.05 | 6.6 |
| EES 66 | 4.0 | 4.5 | ≲0.1 | 5.3 |
| EES 116 | 2.9 | 3.3 | ≲0.1 | 7.8 |
| P 11 | ~1.3 | ≲0.8 | ≲0.2 | ~4.0 |
| *r* 114 | ≲0.2 | ≲0.1 | ≲0.1 | 0.7 |
| *r* 131 | ≲0.06 | ≲0.1 | ≲0.1 | 0.4 |
| *r* 117 | ≲0.05 | ≲0.1 | ≲0.1 | 0.6 |
| *r* 194 | ≲0.04 | ≲0.1 | — | 0.5 |
| *r* 178 | ≲0.01 | ≲0.01 | ≲0.01 | 0.02 |
| *r* 123 | ≲0.01 | ≲0.01 | — | 0.02 |
| *r* 289 | ≲0.01 | ≲0.01 | — | 0.1 |
| EES 122 | ≲0.3 | ≲0.2 | — | 1.1 |
| EES 64 | ≲0.1 | ≲0.2 | — | 1.3 |
| EES 32 | ≲0.02 | ≲0.2 | — | 0.3 |
| P 12 | ≲0.1 | — | — | 0.5 |
| P 5 | ≲0.05 | ≲0.05 | — | 0.9 |
| P 14 | ≲0.04 | — | — | 0.1 |
| P 8 | ≲0.03 | ≲0.01 | — | 0.02 |
| P 13 | ≲0.03 | — | — | 0.3 |
| P 33 | ≲0.03 | — | — | 0.2 |
| P 3 | ≲0.02 | ≲0.02 | — | 0.02 |
| P 4 | ≲0.01 | ≲0.02 | — | 0.08 |
| P 36 | ≲0.01 | — | — | 0.2 |

---

[a] From E.B. Freese (1961)

$\alpha/\beta$, rate of increase of the frequency of revertants per viable phage and per lethal hit, in units $10^{-6}$; *r*, spontaneous mutants; EES, ethyl ethanesulfonate-induced mutants; P, proflavin-induced mutants; —, not measured; ≲, smaller than or about equal to.

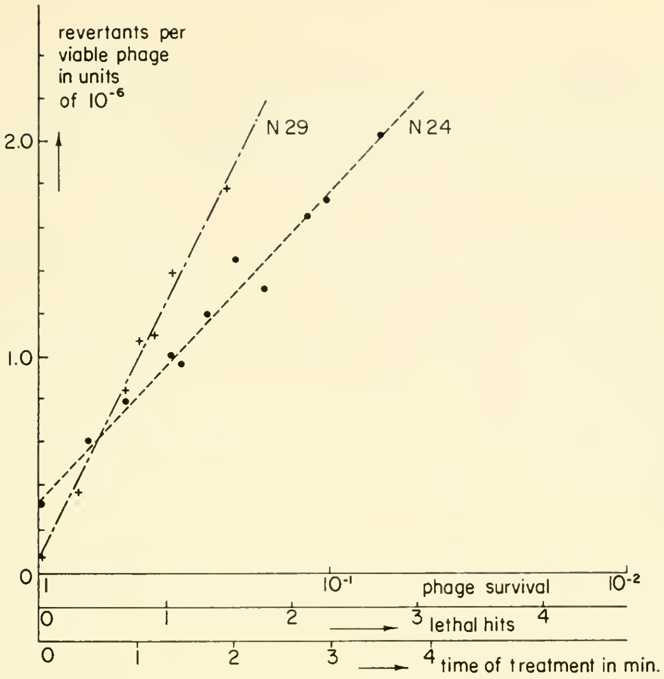


FIG. 17A. The induction of reverse mutations. Reversion induction of BU induced mutants by nitrous acid (see E. B. Freese and E. Freese, 1961).

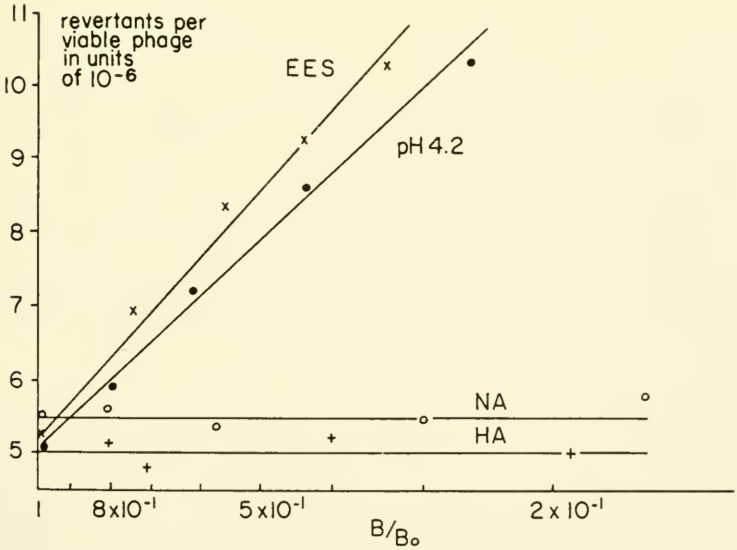


FIG. 17B. The induction of reverse mutations. Reversion induction of an EES-induced mutant (66) by EES and low pH. This mutant does not respond to nitrous acid (NA) or to hydroxylamine (HA) (see E. B. Freese, 1961).

should also be able to induce transversions. Whether they induce deletions is unknown.

The non-transition mutants were therefore tested for their reversion inducibility by EES or low pH. The results of such tests are summarized in Table IV (E. B. Freese, 1961). The values $\alpha/\beta$ in this table are the slopes of the straight line shown in Fig. 17. Each slope has been determined by several points. Several spontaneous and EES induced mutants show a small increase in the frequency of revertants per viable phage, with both low pH and EES, while no such increase was observed with nitrous acid. Several induced revertants of the two EES-induced mutants, EES 66 and EES 116, were examined further and behaved in functional tests as well as in backcrosses as if they arose by back mutations (E. B. Freese, 1962). It thus seems likely that the non-transition mutants which are inducible to revert by low pH and EES have a transversion at their mutant site.

The strongly responding mutants of Table IV probably carry a G-C pair at their mutant site, because EES and low pH seem to remove G much more frequently than A under the conditions used for the induction of mutations. The G-C pair could stem either from a T-A or a C-G pair in the standard DNA. But whether the less strongly responding mutants also have a G-C pair and just react less because of some higher mutagenic specificity or whether they have an A-T pair is not clear.

Some spontaneous and most of the proflavin-induced mutants cannot be induced to revert by EES or low pH. This is complemented by the observation of Orgel and Brenner (1961) that most proflavin- or 5-aminoacridine-induced and some spontaneous mutants can be induced to revert by proflavin or 5-aminoacridine, while BU induced mutants cannot be induced by these agents to revert. If the hypothesis of Brenner *et al.* (1961) is correct, mutants inducible by acridine dyes should contain a deletion or insertion of one nucleotide pair. So far the experimental results do not permit a decision between this possibility and the alternative that these mutants contain a transversion with an A-T pair.

### 4. *The Preferred Mutagenic Direction of Transition Inducers*

Some transition-inducing mutagens react preferentially or predominantly with one of the DNA bases. It should be possible to utilize these agents in order to subdivide transition mutants further according to A-T and G-C pairs at the mutant site. Thus nearly all forward mutants induced by HA and most of the transition mutants induced by EES or low pH should contain a base pair change $\begin{matrix}G\\C\end{matrix} \rightarrow \begin{matrix}A\\T\end{matrix}$.

The directional effect of these mutagens can be observed when the rates of induced reversion are determined for many transition mutants. If one obtains a bipartition into highly and weakly inducible mutants and if the results with different mutagens agree with the expectations, it seems likely that the hypothesis about the direction of base pair changes is correct.

It should be remembered, however, that the frequency of revertants is determined by a selective technique. The mutagenic chemical alters only one base in a pair, and the resulting hybrid may be able to act functionally like a revertant either immediately or only after one or two DNA duplications. Hence the result of measurements on induced reversions depends on the number of DNA duplications which the selective method allows after the first base has been altered. This functional expression may depend, furthermore, on the particular mutant (which of the two DNA strands is attacked, the functional or the other one) and on the mutagen (how much the altered base differs functionally from a normal one). The rate of reversion induction has, therefore, been measured in two ways, at least for nitrous acid (E. B. Freese and E. Freese, 1961) and hydroxylamine (Freese *et al.*, 1961b). On the one hand, the chemically treated T4-*rII* phages were plated directly on the selective bacteria K (rate $\alpha/\beta$). On the other hand, they were preadsorbed onto (UV-killed) bacteria B in which all live phages can replicate about five times; the infected bacteria were then plated on the selective bacteria K (rate $\alpha'/\beta$). The results in Table V show that for HA plating with preadsorption merely enhances the values found for direct plating, while for nitrous acid some rates, which were smaller than others for direct plating, were much larger after preadsorption. The cause of these differences has yet to be resolved.

The rates $\alpha/\beta$ are quite different when the same mutant is induced to revert by different mutagens. This reflects the varying degree of the mutagenic versus the lethal effect. The lethal effect is especially small for HA. If the lethal effect is large, as for EES, the difference between highly and weakly inducible mutants becomes much less pronounced.

Different mutants show greatly different rates of reversion inducibility by a given mutagen. With HA as mutagen a clear bipartition has been observed into highly and weakly (or not at all) inducible mutants; the $\alpha/\beta$ values of these two classes differ by more than a factor 100 (see Table V). It should be mentioned that one more AP mutant (AP 12) has been found highly inducible but has not been included in Table V for reasons given later. The same AP mutants that are most inducible by HA are also most inducible by all other agents used (though for nitrous acid only after preadsorption). But with these other agents the bipartition is not as clear-cut. It is improbable that the higher mutagenic specificity can be responsible for the observed bipartition since HA (and nitrous acid) attack C while low pH and EES attack G and yet the results agree. Thus, most of the rarely or non-inducible mutants have an A-T pair at their mutant site since they have been induced by

TABLE V

Rates of Reversion Induction for Base Analog Inducible Mutants

| Mutant No. | With HA $\alpha/\beta$ | With HA $\alpha'/\beta'$ | With nitrous acid $\alpha/\beta$ | With nitrous acid $\alpha'/\beta'$ | With pH 4.2 $\alpha/\beta$ | With EES $\alpha/\beta$ | Background of spontaneous revertants ($10^{-6}$) |
|---|---|---|---|---|---|---|---|
| AP 114 | 450 | ~1040 | 0.03 | 160 | 8.0 | 1.4 | 5 |
| AP 275 | 280 | 240 | 0.24 | 40 | 7.6 | 1.5 | 2 |
| AP 72 | 134 | 158 | 0.03 | 80 | 5.2 | 1.0 | 1 |
| AP 41 | ~ 62 | ~ 265 | ~0.03 | ~ 28 | ~1.2 | ~2.3 | ~0.3 |
| N 24 | 0.2 | ≲10 | 0.64 | ≲4 | ≲0.1 | ≲0.01 | 0.48 |
| N 29 | 0.1 | ≲ 1 | 1.2 | ≲5 | 0.01 | 0.03 | 0.013 |
| HA 54 | 0.1 | ≲ 3 | — | — | — | — | 0.08 |
| AP 61 | 0.08 | ≲18 | 0.33 | ≲4 | ≲0.03 | ≲0.01 | 0.13 |
| N 19 | 0.06 | ≲ 6 | 0.16 | ≲6 | ≲0.01 | ≲0.02 | 0.009 |
| HA 45 | 0.06 | ≲ 4 | — | — | — | — | 0.01 |
| N 7 | 0.04 | ≲ 2 | 0.15 | ≲8 | ≲0.01 | ≲0.01 | 0.02 |
| HA 11 | 0.04 | ≲ 8 | — | — | — | — | 0.02 |
| N 12 | 0.03 | ≲ 2 | 0.21 | ≲7 | — | ≲0.01 | 0.03 |
| N 17 | 0.03 | — | 0.15 | — | ≲0.01 | ≲0.01 | 0.07 |
| N 31 | 0.03 | ≲ 6 | 0.24 | — |  | ≲0.01 | 0.028 |
| HA 25 | ≲0.001 | ≲ 0.1 | — | — | — | — | 0.003 |
| AP 156 | ≲0.002 | ≲ 5 | 0.03 | ≲4 | 0.01 | 0.08 | 0.007 |
| AP 70 | ≲0.01 | ≲ 6 | 0.01 | ≲1 | 0.01 | 0.07 | 0.02 |
| AP 83 | ≲0.005 | — | 0.07 | ≲2 | 0.01 | 0.09 | 0.01 |

[a] From Freese (1961)
$\alpha/\beta$, rate of increase of the frequency of revertants per viable phage and per lethal hit, in units $10^{-6}$, measured by directly plating phages on bacteria, E. coli B (all viable phages give plaques) and on K (selective, *rII* mutants do not, but revertants do, grow). $\alpha/\beta$ is the same rate but the treated phages were preadsorbed to UV killed bacteria B (survival $10^{-3}$, and the infectd centers plated on K; AP, HA, and N mutants induced by AP, HA, and BU, respectively; — , not measured; ≲ , smaller than or about equal to.

agents that produce transitions. For this conclusion it is unessential that some of the little inducible mutants show induced suppressor mutations, as for example, AP 70, which gives after EES induction only revertants that form big standard type plaques on bacteria B.

One would like to argue that the highly inducible mutants have a G-C pair at their mutant site and that the above agents can induce their back mutations. Here the problem of suppressor mutations does play an important role in the interpretation. Functional tests (on K and B) of induced revertants have shown that for mutants AP 72 and AP 275 (see Table V) most revertants have standard phenotype; for AP 41 both standard type revertants (on K and B) and others with tiny plaques on K (and plaques on B) have been observed; for AP 114 less than 50% of the revertants gave standard phenotype while the rest gave large fuzzy plaques on bacteria B. The important point is that each of these four mutants does yield at least some induced revertants of standard phenotype. When these revertants were backcrossed to standard type phages the frequency of *r* mutants observed in the progeny was not significantly higher than in the stocks used (usually smaller than 0.1%, Freese *et al.*, 1961b, E. B. Freese, 1962). As outlined before these results make it likely that the revertants of standard phenotype arose by genuine back mutations. The four AP mutants in Table V must then have a G-C pair at their mutant site.

The mutant AP 12, although highly inducible, has been omitted from Table V because a functional test revealed that all induced revertants produce *r* plaques on bacteria B and the same was observed with another mutant of this genetic site. Revertants of AP 12, induced by the above agents, may arise only or mainly by suppressor mutations.

The data in Table V indicate that BU induces transitions from GC into AT more frequently then in the opposite direction; this has already been explained earlier (Section III,B,1). In contrast, AP seems to induce transitions in both directions with comparable frequency. Actually all other transition inducing agents predominantly induce changes $\begin{matrix}G\\C\end{matrix} \rightarrow \begin{matrix}A\\T\end{matrix}$. The observed directional effect of the base analogs complements the earlier observation that AP is more efficient than BUdR in inducing the reversion of BU-induced mutants (Freese, 1959b); the opposite statement concerning AP mutants seems no longer justified. Similar directional effects for induced reversions have been observed in *Salmonella typhimurium* (Margolin and Mukai, 1961) when the frequency of AP-induced revertants was measured on plates with minimal and with suboptimally enriched medium. Again, kinetic studies, on the time and mode of appearance of bacterial revertants induced by BUdR or by AP, have been used to identify the mutant base pair in a transition mutant (Rudner, 1961; Strelzoff, 1962).

### 5. *Specificity of Spontaneous Mutations*

The preceding analysis on the induction of mutations also provides information about the specificity of spontaneous mutations. The result will be summarized for the mutations of phage T4:

1. About 80% of the spontaneous mutations are of the non-transition type; the rest apparently arose by transitions (see Table III).
2. Some of the spontaneous non-transition mutants, including some hot spots, seem to carry a transversion (see Table IV). The other non-

transition mutants that are not reversible by EES or low pH may have been caused by transversions, deletions, or insertions.

3. Both transition and transversion mutants which are very susceptible to reversion by agents preferentially attacking G or C also frequently revert spontaneously. It seems likely, therefore, that G-C pairs mutate spontaneously more frequently than A-T pairs.

It must be emphasized that these tentative conclusions are only valid for the particular phage system examined. There is no justification for a general statement about the frequency or the specificity of spontaneous mutations in all organisms. One could argue that G and C are chemically more vulnerable than A or T; but even that is true only for most of the directly acting chemical agents and wrong for certain ones (e.g., peroxides, hydrazine). Besides, we do not know whether in a particular organism most mutations are caused by mistakes during DNA duplication or by chemical changes of DNA. A typical example for the differences of similar organisms has been reported by Kirchner (1960) for *Salmonella;* he found that for one strain (LT2) 26% of the spontaneous mutants could be induced to revert by AP or BUdR while for another strain (LT7) 95% of the spontaneous mutants could be induced to revert by these base analogs. The second strain differs from the first one by a gene which increases the frequency of spontaneous mutations. The result indicates that the mutator gene produces a substance that induces transitions.

## REFERENCES

Alexander, P. (1952). *Nature* **169**, 226.

Alexander, P., and Stacey, K. A. (1958). *Ann. N.Y. Acad. Sci.* **68**, 1225.

Alper, H., and Ebert, M. (1954). *Science* **120**, 608.

Andrew, L. E. (1959). *Am. Naturalist* **93**, 135.

Auerbach, C. (1951). *Cold Spring Harbor Symposia Quant. Biol.* **16**, 199.

Auerbach, C., and Moser, H. (1950). *Nature* **166**, 1019.

Auerbach, C., and Moser, H. (1953). *Z. Induktive Abstammungs-u. Vererbungslehre* **85**, 479, 547.

Auerbach, C., and Robson, J. M. (1946). *Nature* **157**, 302.

Baker, W. K. (1955). *Brookhaven Symposia in Biol.* **8**, 191.

Baron, F., and Brown, D. M. (1955). *J. Chem. Soc.* p. 2855.

Bauch, R. (1948). *Planta* **35**, 536.

Bautz, E., and Freese, E. (1960). *Proc. Natl. Acad. Sci. U.S.* **46**, 1585.

Bayley, C. R., Brammer, K. W., and Jones, A. S. (1961). *J. Chem. Soc.* p. 1903.

Beers, R. F., Hendley, D. D., and Steiner, R. F. (1958). *Nature* **182**, 242.

Benzer, S. (1955). *Proc. Natl. Acad. Sci. U.S.* **41**, 344.

Benzer, S. (1957). *In* "The Chemical Basis of Heredity" (W. D. McElroy and B. Glass, eds.), p. 70. Johns Hopkins Press, Baltimore.

Benzer, S. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 403.

Benzer, S., and Freese, E. (1958). *Proc. Natl. Acad. Sci. U.S.* **44**, 112.

Beukers, R., and Berends, W. (1960). *Biochem. et Biophys. Acta* **49**, 181.

Bradley, D. F., and Wolf, M. K. (1959). *Proc. Natl. Acad. Sci. U.S.* **45**, 944.

Brenner, S., and Smith, J. D. (1959). *Virology* **8**, 124.

Brenner, S., Benzer, S., and Barnett, L. (1958). *Nature* **182**, 983.

Brenner, S., Barnett, L., Crick, F. H. C., and Orgel, A. (1961). *J. Mol. Biol.* **3**, 121.

Brookes, P., and Lawley, P. D. (1960a). *J. Chem. Soc.* p. 539.

Brookes, P., and Lawley, P. D. (1960b). *Biochem. J.* **77**, 478.

Brown, D. M., and Schell, P. (1961). *J. Mol. Biol.* **3**, 709.

Cartledge, J. L., and Blakeslee, A. F. 1934. *Proc. Natl. Acad. Sci. U.S.* **20**, 103.

Chase, M., and Doermann, A. H. (1958). *Genetics* **43**, 332.

Clifton, C. E. (1931). *Proc. Soc. Exptl. Biol. Med.* **28**, 745.

Coughlin, C. A., and Adelberg, E. A. (1956). *Nature* **178**, 531.

Crick, F. H. C., Barnett, L., Brenner, S., and Watts-Tobin, R. J. (1961). *Nature* **192**, 1227.

D'Amato, F. (1950). *Caryologia* **2**, 229.

Darwin, C. (1859). "The Origin of Species by Means of Natural Selection."

De Giovanni, R. (1960). *Ann. N.Y. Acad. Sci.* **84**, Art. 16.

De Giovanni, R. (1961). *Z. Vererbungslehre* **92**, 389.

De Mars, R. I. (1953). *Nature* **172**, 964.

De Mars, R. I., Luria, S. E., Fisher, H., and Levinthal, C. (1953). *Ann. inst. Pasteur* **84**, 113.

Demerec, M. (1937). *Genetics* **22**, 469.

Demerec, M. (1953). *Symposia Soc. Exptl. Biol.* **7**, 44.

Demerec, M. (1960). *Proc. Natl. Acad. Sci. U.S.* **46**, 1075.

Demerec, M., and Hanson, J. (1951). *Cold Spring Harbor Symposia Quant. Biol.* **16**, 215.

Demerec, M., and Hartmann, P. E. (1959). *Ann. Rev. Microbiol.* **13**, 377.

Demerec, M., and Sams, J. (1960). *Proc. Symp. Venice, 1959,* p. 283. *Intern. J. Radiation Biol. Suppl.* p. 283.

De Serres, F. J. (1958). *Genetics* **43**, 187.

De Vries, H. (1909). "The Mutation Theory." Open Court, LaSalle, Illinois.

Deysson, M. (1952). *Compt. rend. acad. sci.* **234**, 650.

Dickey, F. H., Cleland, G. H., and Lotz, C. (1949). *Proc. Natl. Acad. Sci. U.S.* **35**, 581.

Djordjevic, B., and Szybalski, W. (1960). *J. Exptl. Med.* **112**, 509.

Doermann, A. H., Chase, M., and Stahl, F. W. (1955). *J. Cellular Comp. Physiol.* **45** (Suppl. 2), 51.

Doty, P., Marmur, J., and Sneoke, N. (1959). *Brookhaven Symposia in Biol.* **12**, 1.

Dunn, D. B., and Smith, J. D. (1954). *Nature* **174**, 304.

Dunn, D. B., and Smith, J. D. (1955). *Nature* **175**, 336.

Eidinoff, M. L., Cheong, L., and Rich, M. A. (1959). *Science* **129**, 1550.

Ellmauer, H., and Kaplan, R. W., 1959. *Naturwiss.* **46**, 150.

Endo, H. W., Miura, K., Hidaka, F., and Hiruk, C. (1961). *Nature* **190**, 833.

Fahmy, O. G., and Fahmy, M. J. (1955). *J. Genet.* **53**, 563.

Fahmy, O. G., and Fahmy, M. J. (1957). *Nature* **180**, 31.

Fahmy, O. G., and Fahmy, M. J. (1958). Discussion to the paper on the "Mutagenic Effects of Alkylating Agents," presented by C. Auerbach in *Ann. N.Y. Acad. Sci.* **68**, 736.

Fischer, R., and McKinley, E. B. (1927). *J. Infectious Diseases* **40**, 399.

Fosse, R., Hieulle, A., and Bass, L. W. (1924). *Compt. rend. acad. sci.* **178**, 811.

Fraenkel-Conrat, H. (1954). *Biochim. et Biophys. Acta* **15**, 307.
Fraser, D. K. (1957). *Virology* **3**, 527.
Frederiksen, S., and Klenow, H. (1960). *Biochem. Biophys. Research Communs.* **2**, 329.
Freese, E. (1959a). *J. Mol. Biol.* **1**, 87.
Freese, E. (1959b). *Proc. Natl. Acad. Sci. U.S.* **45**, 622.
Freese, E. (1959c). *Brookhaven Symposia in Biol.* **12**, 63.
Freese, E. (1961). *Proc. 5th Intern. Congr. Biochem., Moscow, 1961.*
Freese, E., Freese, E. B., and Bautz, E. (1961a). *J. Mol. Biol.* **3**, 133.
Freese, E., Freese, E. B., and Bautz, E. (1961b). *Proc. Natl. Acad. Sci. U.S.* **47**, 845.
Freese, E. B. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 540.
Freese, E. B. (1962). *Biochem. Biophys. Research Communs.* **7**, 18.
Freese, E. B., and Freese, E. (1961). *Virology* **13**, 19.
Fresco, J. R., and Alberts, B. M. (1960). *Proc. Natl. Acad. Sci. U.S.* **46**, 311.
Fricke, H., and Demerec, M. (1937). *Proc. Natl. Acad. Sci. U.S.* **23**, 320.
Friedewald, W. F., and Anderson, R. S. (1940). *Proc Soc. Exptl. Biol. Med.* **45**, 713.
Gard, S., and Maaløe, O. (1959). *In* "The Viruses" (F. M. Burnet and W. M. Stanley, eds.), Vol. 1, Chapter V, pp. 359–418. Academic Press, New York.
Geiduschek, E. P. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 950.
Giles, N. H. (1955). *Brookhaven Symposia in Biol.* **8**, 103.
Giles, N. H., De Serres, J., and Partridge, C. W. H. (1955). *Ann. N.Y. Acad. Sci.* **59**, 536.
Glaess, E. (1956). *Chromosoma* **8**, 260.
Glass, B. (1955). *Brookhaven Symposia in Biol.* **8**, 148.
Glass, E. A., and Novick, A. (1959). *J. Bacteriol.* **77**, 10.
Gottschling, H., and Freese, E. (1961). *Z. Naturforsch.* **16b**, 515.
Green, D. MacD., and Krieg, D. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 64.
Greer, S. B. (1958). *J. Gen. Microbiol.* **18**, 544.
Greer, S. B. (1960). *J. Gen. Microbiol.* **22**, 618.
Greer, S. B., and Zamenhof, S. (1962). *J. Mol. Biol.* **4**, 123.
Gutz, H. (1961). *Nature* **191**, 1124.
Haas, F. L., and Doudney, C. O. (1957). *Proc. Natl. Acad. Sci. U.S.* **43**, 871.
Hadorn, E., and Niggli, H. (1946). *Nature* **157**, 162.
Hakala, M. T. (1959). *Biol. Chem.* **234**, 3072.
Handschumacher, R. E., and Welch, A. D. (1960). *In* "The Nucleic Acids" (E. Chargaff and J. N. Davidson, eds.), Vol. 3, pp. 453–526. Academic Press, New York.
Harm, W. (1960). *Z. Vererbungslehre* **91**, 52.
Heiner, R. E., Konzak, C. F., Nilan, R. A., and Legault, R. R. (1960). *Proc. Natl. Acad. Sci. U.S.* **46**, 1215.
Helinski, D. R., and Yanofsky, C. (1961). *Proc. Natl. Acad. Sci. U.S.* **48**, 173.
Hershey, A. D., and Chase, M. (1952). *J. Gen. Physiol.* **36**, 39.
Hiatt, C. W. (1960). *Trans. N.Y. Acad. Sci.* **23**, 66.
Hirota, Y. (1960). *Proc. Natl. Acad. Sci. U.S.* **46**, 57.
Hirota, Y., and Tijima, T. (1957). *Nature* **180**, 655.
Hollaender, A., and Kimball, R. F. (1956). *Nature* **177**, 726.
Iyer, V. N., and Szybalski, W. (1958). *Proc. Natl. Acad. Sci. U.S.* **44**, 446.
Iyer, V. N., and Szybalski, W. (1959). *Science* **129**, 839.
Jacob, F. (1954). *Compt. rend. acad. sci.* **238**, 732.
Jensen, K. A., Kolmark, G., and Westergaard, M. (1949). *Hereditas* **35**, 521.

Jensen, K. A., Kirk, J., Kolmark, G., and Westergaard, M. (1951). *Cold Spring Harbor Symposia Quant. Biol.* **16**, 245.

Kanazir, D. (1958). *Biochem. et Biophys. Acta* **30**, 20.

Kaplan, M. S., and Tomlin, P. A. (1960). *Radiation Research* **12**, 447.

Kaudewitz, F. (1959). *Nature* **83**, 1829; *Z. Naturforsch.* **146**, 528.

Kaufman, E., and Hiatt, C. W. (1959). *Virology* **9**, 478.

Kihlman, B. (1955). *Exptl. Cell Research* **8**, 345.

Kihlman, B. (1960). *Exptl. Cell Research* **20**, 657.

Kirby-Smith, J. S., and Craig, D. L. (1955). *Records Gen. Soc. Am.* **24**, 579.

Kirchner, C. W. J. (1960). *J. Mol. Biol.* **2**, 331.

Koch, A. L., and Lamond, W. A. (1956). *J. Biol. Chem.* **219**, 189.

Kolmark, G., and Giles, N. H. (1955). *Genetics* **40**, 890.

Kolmark, G., and Westergaard, M. (1953). *Hereditas* **39**, 209.

Konrad, M. (1960). *Ann. N.Y. Acad. Sci.* **84**, 678.

Kornberg, S. R., Zimmerman, S. B., and Kornberg, A. (1961). *J. Biol. Chem.* **236**, 1487.

Kozinski, A. W., and Szybalski, W. (1959). *Virology* **9**, 260.

Krieg, D. R. (1959a). *Virology* **8**, 80.

Krieg, D. R. (1959b). *Virology* **9**, 215.

Lamarck, J. B. (1809). "Philosophie zoologique."

Latarjet, R. (1949). *Compt. rend. acad. sci.* **228**, 1354.

Latarjet, R. (1954). *Acta Unio Intern. contra Cancrum* **10**, 136.

Lea, D. E. (1955). "Action of Radiations on Living Cells," 2nd ed. Cambridge Univ. Press, London and New York.

Lehman, J. R., and Pratt, D. (1960). *J. Biol. Chem.* **235**, 3254.

Lerman, L. S. (1961). *J. Mol. Biol.* **3**, 18.

Lett, J. T., Alexander, P., and Parkins, G. M. (1962). *Arch. Biochem. Biophys.* **97**, 80.

Leuppold, V. (1962). *Arch. Julius-Klaus-Stift. Vererbungsforsch.* In preparation.

Levan, A. (1951). *Cold Spring Harbor Symposia Quant. Biol.* **16**, 233.

Litman, R. M., and Ephrussi-Taylor, H. (1959). *Compt. rend. acad. sci.* **249**, 838.

Litman, R. M., and Pardee, A. B. (1956). *Nature* **178**, 529.

Litman, R. M., and Pardee, A. B. (1959). *Virology* **8**, 125.

Litman, R. M., and Pardee, A. B. (1960). *Biochem. et Biophys. Acta* **42**, 117, 131.

Littlefield, J. W., and Dunn, D. B. (1958). *Biochem. J.* **70**, 642.

Lorkiewicz, Z., and Szybalski, W. (1960). *Biochem. Biophys. Research Communs.* **2**, 413.

Lorkiewicz, Z., and Szybalski, W. (1961). *J. Bacteriol.* **82**, 195.

Loveless, A. (1951). *Nature* **167**, 338.

Loveless, A. (1959). *Proc. Roy. Soc.* **B150**, 486 and 497.

Luria, S. E. (1947). *Proc. Natl. Acad. Sci. U.S.* **33**, 253.

Luria, S. E., and Delbrück, M. (1943). *Genetics* **28**, 491.

McKinley, E. B., Fischer, R., and Holden, M. (1926). *Proc. Soc. Exptl. Biol. Med.* **23**, 408.

Margolin, P., and Mukai, F. H. (1961). *Z. Vererbungslehre* **92**, 330.

Marinelli, L. D., and Taylor, L. S. (1954). *Radiation Biol.* **1**, 145.

Marmur, J., Schildkraut, C. L., and Doty, P. (1961). *J. de Chimie Physique* p. 945.

Marquardt, H. (1938). *Z. Botan.* **32**, 401.

Mashima, S., and Jkeda, Y. (1958). *J. Appl. Microbiol.* **6**, 45.

Mendel, G. (1865). *Verhandl. naturforsch. Ver. Brünn, Abhandl.* **VI**.

Meselson, M., and Stahl, F. W. (1958). *Proc. Natl. Acad. Sci. U.S.* **44**, 671.

Miles, H. T. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 791.

Miyake, T. (1959). *Genetics* **45**, 1, 11.

Muller, H. J. (1927). *Science* **66**, 382.

Mundry, K. W., and Gierer, A. (1958). *Z. Vererbungslehre* **89**, 614.

Nasrat, G. E., Kaplan, W. D., and Auerbach, C. (1954). *Z. Induktive Abstammungs-u. Vererbungslehre* **86**, 249.

Novick, A. (1956). *Brookhaven Symposia in Biol.* **8**, 201.

Oehlkers, F. (1943). *Z. Induktive Abstammungs-u. Vererbungslehre* **81**, 313.

Oehlkers, F. (1949). *Heidelberger Akad. Wiss.* **9**, 373.

Oehlkers, F. (1953). *Heredity* **6** (Suppl.), 95.

Orgel, A., and Brenner, S. (1961). *J. Mol. Biol.* **3**, 762.

Plough, H. H. (1941). *Cold Spring Harbor Symposia Quant. Biol.* **9**, 127.

Pomper, S., and Atwood, K. C. (1955). *Radiation Biol.* **2**, 431.

Pratt, D., and Stent, G. S. (1959). *Proc. Natl. Acad. Sci. U.S.* **45**, 1507.

Price, C. C. (1958). *Ann. N.Y. Acad. Sci.* **68**, 663.

Putnam, F. W. (1953). *In* "The Proteins" (H. Nenrath and K. Bailey, eds.), Vol. I, Part B, p. 893. Academic Press, New York.

*Radiation Biol.* (1954, 1955).

*Radiation Research* (1959). Suppl. 1.

Rapoport, J. A. (1946). *Doklady Akad. Nauk. S.S.S.R.* **54**, 65.

Rapoport, J. A. (1947a). *Doklady Akad. Nauk. S.S.S.R.* **56**, 12.

Rapoport, J. A. (1947). *Doklady Vsesoyuz. Akad. Sel'sko-khoz Naukim. V. I. Lenina* **12**, 12.

Rapoport, J. A. (1948a). *Doklady Akad. Nauk. S.S.S.R.* **59**, 1183.

Rapoport, J. A. (1948b). *Doklady Akad. Nauk. S.S.S.R.* **60**, 469.

Reiner, B., and Zamenhof, S. (1957). *J. Biol. Chem.* **228**, 475.

Rogers, S. (1957). *J. Exptl. Med.* **105**, 279.

Ross, W. C. J. (1958). *Ann. N.Y. Acad. Sci.* **68**, 669.

Rudner, R. (1960). *Biochem. Biophys. Research Communs.* **3**, 275.

Rudner, R. (1961). *Z. Vererbungslehre* **92**, 336 and 361.

Ryan, F. J. (1959). *J. Gen. Microbiol.* **21**, (No. 3), 530.

Ryan, F. J. (1962). *In* "Methodology in Basic Genetics" (W. J. Burdette, ed.), Holden-Day, Inc., San Francisco.

Ryan, F. J., Nakada, D., and Schneider, M. J. (1961). *Z. Vererbungslehre* **92**, 38.

Sauerbier, W. (1960). *Nature* **188**, 327.

Sax, K. (1937). *Genetics* **22**, 523.

Sax, K. (1941). *Cold Spring Harbor Symposia Quant. Biol.* **9**, 93.

Sax, K. (1957). *Quart. Rev. Biol.* **32**, 15.

Schuster, H. (1960). *Biochem. Biophys. Research Communs.* **2**, 320.

Schuster, H. (1961). *J. Mol. Biol.* **3**, 447.

Schuster, H., and Schramm, G. (1958). *Z. Naturforsch.* **136**, 697.

Shugar, D. (1960). *In* "The Nucleic Acids" (E. Chargaff and J. N. Davidson, eds.), Vol. 3, Chapter 30, p. 39. Academic Press, New York.

Sinsheimer, R. L. (1955). *J. Biol. Chem.* **215**, 579.

Sobels, F. H. (1956). *Nature* **177**, 979.

Spencer, M. (1959). *Acta Cryst.* **12**, 59, 66.

Stacey, K. A., Cobb, M., Cousens, S. F., and Alexander, P. (1958). *Ann. N.Y. Acad. Sci.* **68**, 682.

Stadler, L. J. (1928). *Science* **68**, 186.

Staehlin, M. (1958). *Biochem. et Biophys. Acta* **29**, 410.

Stahl, F. W. (1959). *In* "The Viruses" (F. M. Burnett and W. M. Stanley, eds.), Vol. 2, Chapter X, pp. 353–384. Academic Press, New York.

Stahl, F. W., Crasemann, J. M., Okun, L., Fox, E., and Laird, C. (1961). *Virology* **13**, 98.

Steiner, R. F., and Beers, R. F. (1958). *Science* **127**, 335.

Stent, G. S. (1958). *Advances in Virus Research* **5**, 95.

Stevens, M. A., Magrath, D. J., Smith, H. W., and Brown, G. B. (1958). *J. Biol. Chem.* **233**, 1513.

Stone, W. S. (1955). *Brookhaven Symposia in Biol.* **8**, 171.

Stone, W. S., Wyss, O., and Haas, F. (1947). *Proc. Natl. Acad. Sci. U.S.* **33**, 59.

Strauss, B. S. (1962). *J. Bacteriol.* **83**, 241.

Strauss, B. S., and Okubo, S. (1960). *J. Bacteriol.* **79**, 464.

Strelzoff, E. (1962). Ph. D. Thesis, Columbia University, New York.

Swanson, C. P., and Stadler, L. J. (1955). *Radiation Biol.* **2**, 249.

Szybalski, W. (1958). *Ann. N.Y. Acad. Sci.* **76**, 475.

Szybalski, W., and Djordjevic, B. (1959). *Genetics* **44**, 540.

Szybalski, W., Opara-Kubinska, Z., Lorkiewicz, Z., Ephrati-Elizur, E., and Zamenhof, S. (1960). *Nature* **188**, 743.

Takemura, S. (1957). *J. Biochem.* **44**, 321.

Takemura, S. (1959). *Bull. Chem. Soc. Japan* **32**, 920.

Tamm, C., Hodes, M. E., and Chargaff, E. (1952). *J. Biol. Chem.* **195**, 49.

Tamm, C., Shapiro, H. S., Lipschitz, R., and Chargaff, E. (1953). *J. Biol. Chem.* **203**, 673.

Tessman, E. S. (1956). *Virology* **2**, 679.

Tessman, I. (1959). *Virology* **9**, 375.

Thoday, J. M., and Read, J. (1947). *Nature* **160**, 608.

Timofeeff-Ressovsky, M. W., Zimmer, K. G., and Delbrück, M. (1935). *Nachr. Ges. Wiss. Göttingen Math. Physik. Kl. Fachgruppen* [VI] **1**, 189.

Treffers, H. P., Spinelli, V., and Belser, N. D. (1954). *Proc. Natl. Acad. Sci. U.S.* **40**, 1064.

Vielmetter, W., and Schuster, H. (1960). *Biochem. Biophys. Research Communs.* **2**, 324.

Vielmetter, W., and Wieder, C. M. (1959). *Z. Naturforsch.* **14b**, 312.

Vogt, M. (1950). *Z. Induktive Abstammungs-u. Vererbungslehre* **83**, 324.

Wacker, A., Kirschfeld, S., and Trager, L. (1960a). *J. Mol. Biol.* **2**, 241.

Wacker, A., Dellweg, H., and Weinblum, D. (1960b). *Naturwiss.* **47**, 477.

Watson, J. D., and Crick, F. H. C. (1953). *Cold Spring Harbor Symposia Quant. Biol.* **18**, 123.

Weigle, J. J. (1953). *Proc. Natl. Acad. Sci. U.S.* **39**, 628.

Weigle, J. J., and Dulbecco, R. (1953). *Experientia* **9**, 372.

Westergaard, M. (1957). *Experientia* **13**, 224.

Weygand, F., Wacker, A., and Dellweg, H. (1952). *Z. Naturforsch.* **7b**, 19.

Wilkins, M. H. F. (1956). *Cold Spring Harbor Symposia Quant. Biol.* **21**, 75.

Witkin, E. M. (1947). *Cold Spring Harbor Symposia Quant. Biol.* **12**, 256.

Witkin, E. M. (1956). *Cold Spring Harbor Symposia Quant. Biol.* **21**, 123.

Witkin, E. M. (1959). *Proc. 10th Intern. Congr. Genet., Montreal, 1958*, p. 280.

Wolff, S. (1959). *Radiation Research Suppl.* **1**, 453.

Wolff, S., and Luippold, H. E. (1955). *Science* **122**, 231.

Wyss, O., Clark, J. B., Haas, F., and Stone, W. S. (1948). *J. Bacteriol.* **56**, 51.

Yamamoto, N. (1956). *Virus (Osaka)* **6**, 510.

Yamamoto, N. (1958). *J. Bacteriol.* **75**, 443.

Zamenhof, S., and Greer, S. (1958). *Nature* **182**, 611.

Zamenhof, S., and Griboff, G. (1954). *Nature* **174**, 306.

Zetterberg, G. (1960). *Exptl. Cell Research* **20**, 659.

## *Chapter* VI

# Biosynthesis of RNA in Relation to Genetic Coding Problems

ELLIOT VOLKIN

*Biology Division,*
*Oak Ridge National Laboratory,*[1]
*Oak Ridge, Tennessee*

I. Introduction . . . . . 271
II. Evidence Implicating an RNA Species as a Carrier of Genetic Information . . . . . 273
  A. The DNA-like RNA in Bacteriophage Infection . . . . . 273
  B. Gradient Centrifugation of Cell Components . . . . . 273
  C. Messenger RNA-Homologous DNA Hybrids . . . . . 275
  D. RNA Turnover . . . . . 275
  E. Messenger RNA in Other Biologic Systems . . . . . 276
III. Mechanism of Messenger RNA Function Based on *in Vivo* Experiments . . . . . 277
  A. Enzyme Induction . . . . . 277
  B. Bacteriophage Infection . . . . . 278
IV. Attempts at Isolation of a Messenger RNA . . . . . 279
V. *In Vitro* Synthesis of RNA . . . . . 280
  A. Polynucleotide Phosphorylase . . . . . 280
  B. RNA Polymerase . . . . . 280
VI. RNA Involvement in *in Vitro* Protein Synthesis . . . . . 281
  A. Requirement for DNA-Directed RNA Synthesis . . . . . 281
  B. The Use of Polynucleotides in Establishing the Code . . . . . 282
  C. Correlations with Plant Virus Mutants . . . . . 284
VII. The Triplet Code . . . . . 284
VIII. Summary . . . . . 286
  References . . . . . 286

## I. INTRODUCTION

A tremendous concentration of research effort on the nucleic acids from the late 1940's through the early 1960's has culminated in what many respected investigators believe to be the essentially complete answer to molecular problems of inheritance and specific growth. Since their discovery by Miescher in 1871, the nucleic acids have been variously implicated as occupying a central role in the processes of genetic continuity, though historically such conjecture has been based on evidence of circumstance in time and place (Brachet, 1933, 1937, 1940a; Caspersson, 1941). As with so many scientific endeavors, research in the nucleic acids acquired a renewed impetus with breakthroughs in methodology, specifically the development of chromatographic procedures (Vischer and Chargaff, 1947, 1948a,b; Cohn, 1949, 1950a,b, 1951) and the availability of radioactive elements. The establishment of deoxyribonucleic acid (DNA) as the entity responsible for cellular heritability has been thoroughly documented (see Chapters I, II, and IV). This chapter will be largely concerned with one aspect of ribonucleic acid (RNA) metabolism, i.e., the role of RNA in the translation of genetic information from DNA to the sites of protein synthesis.

---

[1] Operated by Union Carbide Corporation for the United States Atomic Energy Commission.

On the basis of rigorous and detailed *in vivo* and *in vitro* experimentation, originally done by Zamecnik and colleagues (see review by Zamecnik, 1960), such sites for protein synthesis have been demonstrated to be the RNA-rich subcellular components of cells—the microsomes of highly developed tissues, the ribosomes of microbial systems (see Chapter VIII). Although the function of the bulk of RNA in these particles is still open to question, the role of another species of RNA has been disclosed by the use of cell-free systems that synthesize protein. This is the "transfer RNA," which is involved in carrying amino acids in their aminoacyl form to the particles where they are utilized for protein synthesis. Purification and separation of transfer RNA's have revealed that a high degree of specificity exists between the transfer RNA species and particular amino acids (see Chapter VIII). This phase of specific interaction, however, is not considered to occupy the key role in determining the ultimate explicit structure of the protein product. For this purpose, another species of RNA has been invoked. The generally accepted term for it is "messenger RNA" (Jacob and Monod, 1961), so named because of its putative intermediary role in relaying the genetic code of DNA to functional sites. Within the span of just a few years, research in this area has quickly passed from that which purported to demonstrate the biologic existence of such an RNA to that which now utilizes RNA *in vitro* systems to establish the exact coding relations between specific nucleotide triplets and specific amino acids incorporated into protein. At this point, the speed with which such coding information becomes available in the literature may make this writing archaic by the time it reaches final printing.

## II. EVIDENCE IMPLICATING AN RNA SPECIES AS A CARRIER OF GENETIC INFORMATION

### A. THE DNA-LIKE RNA IN BACTERIOPHAGE INFECTION

In 1956 Volkin and Astrachan observed that in bacteriophage-infected *Escherichia coli* B a minor species of RNA was synthesized whose composition appeared to be similar to the DNA of the phage but not to the host. This biologic system offers a unique advantage for such a study, since viral infection immediately eliminates the bulk of RNA synthesis, and the synthesis of a minor class of RNA having a different over-all composition can be magnified. Thus, it was possible to demonstrate the existence of phage-specific RNA even during T7 infection (Volkin *et al.*, 1958), as well as that produced during T2 infection. The DNA composition of T7 deviates only slightly from that of the host, whereas the DNA composition of T2 is widely different. The demonstration of phage-specific RNA originally took the form of a comparison between relative specific activities of the RNA mononucleotides liberated by alkaline hydrolysis, after an assimilation of $P^{32}$ orthophosphate (Volkin and Astrachan, 1956a,b, 1957; Volkin *et al.*, 1958). The conclusions from these experiments, that a phage DNA-like RNA was being synthesized, has since been amply confirmed by employing a $C^{14}$-labeled precursor (Volkin, 1959), and by a comparative study of the rates of labeling of the free nucleotide pool with the rates of labeling of the RNA mononucleotides (Astrachan, 1960; Volkin, 1962).

By employing only brief isotope assimilation periods (pulse experiments), it was revealed that the newly synthesized RNA does not merely accumulate but undergoes constant breakdown and renewal, or turnover (Volkin and Astrachan, 1957; Astrachan and Volkin, 1958). It was demonstrated (Astrachan and Volkin, 1958) that such turnover, though always present to some extent during the intracellular production of phage, is most pronounced during the first 10–15 minutes after the initial infection. It is noteworthy that virtually all of the extensive biochemical transformations of the host take place during this initial time period (Kornberg *et al.*, 1959; Flaks *et al.*, 1959; Bessman, 1959; Sommerville and Greenberg, 1959; Keck *et al.*, 1960).

### B. GRADIENT CENTRIFUGATION OF CELL COMPONENTS

The question of a specific RNA newly induced by bacteriophage infection has been reinvestigated in a number of other laboratories by different experimental approaches. Their investigations not only reinforce the conclusions of Volkin and Astrachan but additionally describe the subcellular localization of the new RNA, and, in one instance, demonstrate a *bona fide* complexing between phage-specific RNA and the homologous phage DNA.

By employing a sucrose gradient technique for the fractionation of bacterial intracellular particles, Nomura *et al.* (1960) showed that the 30S particles contained most of the radioactive RNA after a short-term labeling period of T2-infected *E. coli* with $P^{32}$. A large part of this radioactive RNA could be released, depending on the $Mg^{++}$ concentration, yielding a nucleic acid essentially free of protein with a sedimentation constant of 8S. Some of the latter exists in the "soluble" form in the T2-infected cell. The T2-specific RNA has a distinguishably higher electrophoretic mobility than normal ribosomal RNA.

Brenner *et al.* (1961) employed the elegant procedure of separating components of differing densities by equilibrium centrifugation in a cesium chloride gradient. They, too, concluded that RNA with a base composition similar to that of phage DNA is associated with ribosomes and can be detached by lowering the magnesium concentration. However, in this case, the new RNA is associated with 70S and 100S ribosomes, and when disassociated has a sedimentation constant of 12S. In accord with the findings of Volkin and Astrachan, their data clearly demonstrate a turnover of the newly synthesized RNA, whether labeling was carried out with $C^{14}$-uracil or $P^{32}O_4$. They further demonstrated that no new ribosomes are formed after T2 infection, but that these particles become enriched for T2-specific or T2 messenger RNA. Similar experiments employing $S^{35}$ for protein labeling revealed that the ribosomes are also enriched for nascent protein that undergoes rapid turnover with respect to its association to the ribosomes. These data led them to the conclusion that pre-existing ribosomes synthesize most of the protein in the infected cells but under the complete direction of the messenger RNA they contain, uninfluenced by the bulk of stable ribosomal RNA. Gros *et al.* (1961), using sucrose gradient separation techniques, concluded that normal, actively growing *E. coli* contains a messenger RNA component, physically distinct from the bulk ribosomal RNA and transfer RNA. They find the active RNA associated with 70S ribosomes, which Tissières *et al.* (1960) have described as the principal site of *in vitro* protein synthesis. Turnover of the messenger RNA in uninfected *E. coli* was clearly demonstrated; in this case, however, turnover consists of a transfer of newly labeled RNA to metabolically stable ribosomal RNA and soluble RNA, rather than a breakdown of RNA to acid-soluble products.

### C. MESSENGER RNA-HOMOLOGOUS DNA HYBRIDS

Perhaps the most convincing evidence for the existence of a phage-specific messenger RNA has been accomplished by Hall and Spiegelman (1961). These workers demonstrated that when single-stranded T2 DNA was slowly cooled in the presence of partially purified T2-specific RNA, an RNA-DNA hybrid was formed that could be separated by equilibrium density centrifugation. Since such hybrid formation was found to be unique to the homologous pair, the conclusion seems valid that a high degree of homology exists between T2 DNA and T2-specific RNA in terms of their complementary nucleotide sequential patterns. A later study by Spiegelman *et al.* (1961) demonstrated that such DNA-RNA complexes could be detected as naturally occurring hybrids in T2-infected *E. coli*. Hayashi and Spiegelman (1961) have extended these observations to bacterial systems that were subjected to a shift from an enriched medium to a basal medium. These systems simulate conditions of phage infection whereby ribosomal RNA synthesis is depressed while protein synthesis and messenger RNA synthesis proceed preferentially. The RNA synthesized and labeled under these "step-down" cultural conditions possessed the ability to form RNA-DNA hybrids when slow-cooled with only their homologous DNA's.

### D. RNA TURNOVER

The question of turnover or metabolic instability of messenger RNA, although referred to later with respect to its functional role, warrants further comment here. The findings of Volkin and Astrachan (Volkin and Astrachan, 1957; Astrachan and Volkin, 1959; Volkin, 1962) with T2-infected *E. coli* demonstrate clearly that the T2-specific RNA is labile and converted to acid-soluble products. Furthermore, these breakdown products do not, to any significant extent, reach the level of nucleosides and inorganic phosphate (Volkin, 1962). The phosphorus originating in messenger RNA can eventually terminate in phage DNA but not by either an obligatory pathway or a pathway utilizing large (acid-insoluble) polynucleotide fragments (Volkin, 1962). The *in vitro* data of Cohen *et al.* (1961) would indicate that the phage-specific RNA is preferentially enzymatically hydrolyzed to ribonucleoside-5′-$PO_4$'s, which are readily reutilized for phage DNA synthesis.

In growing cultures of bacteria, a turnover of messenger RNA can be demonstrated with respect to (*1*) its transient existence evidenced by isotopic labeling, and (*2*) its geographical translocation from one type of ribosomal particle to another. It should be noted, however, that there is uniform and overwhelming evidence that a significant measurable loss of acid-insoluble (or high molecular weight) RNA never occurs in normally growing bacteria. Thus, in pulse-type experiments, after removal of the radioactive substrate, instead of a loss of label, there is a continued increase in radioactivity into total RNA for some time. Concurrently, there is a shift in labeling pattern of the RNA nucleotides in a manner that would indicate an over-all transition of DNA-like RNA to ribosomal, or total, RNA (Volkin and Astrachan, 1957; Hayashi and Spiegelman, 1961). These observations pose the question of the fate of messenger RNA and its probable transformation to stable ribosomal RNA. This aspect of the problem is covered in more detail in Chapter VII. Suffice it to note here that since most organisms and tissues contain ribosomal RNA of high cytidylic and particularly high guanylic acid contents, efficient conversion of messenger RNA to ribosomal RNA could indicate that large amounts of guanylic and cytidylic acids have been added. Kitazume *et al.* (1962) studied the kinetics of synthesis of the dRNA [*sic*] in yeast subjected to differing metabolic conditions. They conclude that dRNA is an obligatory precursor to ribosomal RNA and that this transition occurs by a subtraction of nucleotides from dRNA. They further propose that instability of dRNA is not necessarily a result of its biologic action (see below) but merely results from conversion to ribosomal RNA.

### E. MESSENGER RNA IN OTHER BIOLOGIC SYSTEMS

Yčas and Vincent (1960), comparing the labeling of RNA mononucleotides liberated as the 2′-, 3′-isomers (alkaline hydrolysis) and as the 5′-nucleotides (phosphodiesterase hydrolysis), concluded that in exponentially growing yeast an RNA with a composition like yeast DNA exists. Astrachan and Fisher (1961) and Volkin (1962) used 30-second pulses of $P^{32}$-labeling to demonstrate the existence of a DNA-like RNA in *Proteus vulgaris*, *E. coli*, and *Pseudomonas aeruginosa*—AT/GC ratios of these organisms vary from 1.47 to 0.92 to 0.50, respectively. In such growing bacteria, if the isotope incorporation period is extended to a few minutes, the labeling pattern appears to be identical with total (or ribosomal) RNA, in accord with the observations that messenger RNA is present as a minor fraction of total RNA. The occurrence of a messenger RNA in *E. coli* and *P. aeruginosa* has been independently noted by Hayashi and Spiegelman (1961), who likewise determined its presence in *Bacillus megaterium*, an organism with DNA of high AT/GC ratio. The observations of messenger RNA in growing *E. coli* by Gros *et al.* (1961) have been mentioned.

Some evidence exists for the presence of an RNA with a composition like the homologous DNA in mammalian systems (Boulanger and Montreuil, 1952; Hokin and Hokin, 1954; Logan and Davidson, 1957). Here again, such observations were attainable only where relatively short periods of isotope feeding were employed and the composition of newly formed RNA was an apparent one, evaluated from the specific radioactivity of its mononucleotides. With the advent of autoradiography, particularly where tritium-labeled precursors are involved, experimentation with a large variety of organisms makes it increasingly clear that most, if not all, RNA synthesis of the cell starts within the nucleus —the initial product may be exclusively a DNA-like RNA (see Chapter II).

## III. Mechanism of Messenger RNA Function Based on *in Vivo* Experiments

### A. ENZYME INDUCTION

Perhaps the most striking characteristic of messenger RNA function in microbial systems is its short lifetime with respect to activity. That is, if it behaves either as a template or a catalyst, it must be an extremely unstable one. Evidence by way of independent techniques suggests that inactivation of messenger RNA is a direct result of its action, and, as a matter of fact, a reasonable degree of stoichiometry between messenger RNA and biologic product has been demonstrated.

The synthesis of the enzyme $\beta$-galactosidase is initiated almost immediately at a maximum rate by addition of the inducer, while synthesis is stopped in an equally short time interval by removal of the inducer (Monod, 1958; Boezi and Cowie, 1961; Roberts *et al.*, 1961). Essentially similar kinetics of $\beta$-galactosidase synthesis are obtained by adding and removing the responsible gene (Pardee *et al.*, 1959; Riley *et al.*, 1960). Insertion of the $\beta$-galactosidase gene by bacterial conjugation results in enzyme synthesis within 2 minutes; moreover, the subsequent rate of enzyme synthesis per zygote is constant. Experiments with heavily labeled ($P^{32}$) male parental bacteria, involving ample time for gene expression after insertion, followed by various periods of $P^{32}$ decay, reveal that the rate of enzyme synthesis decreases as a function of $P^{32}$ atoms decayed. Thus the integrity of the $z^+$ gene is constantly required for continued enzyme synthesis. These data on $\beta$-galactosidase induction are interpreted by Jacob and Monod (1961) to mean that the structural messenger of the gene is rapidly formed and exists as a short-lived intermediate. The group at the Carnegie Institution of Washington (Roberts *et al.*, 1961; Boezi *et al.*, 1961) has described similar kinetics for the induction of $\beta$-galactosidase and has suggested a model for the induction mechanism that similarly proposes an unstable enzyme-forming unit that must be synthesized continuously.

Further experiments with inducible enzyme systems utilizing purine and pyrimidine analogs not only confirm the functional instability of the gene messenger but suggest that this intermediate is indeed a polyribonucleotide. The presence of 5-fluorouracil results in an alteration of the $\beta$-galactosidase protein whereby its enzymatic activity is drastically reduced, but its antigenic capacity is retained (Bussard *et al.,* 1960). The same analog, 5-fluorouracil, causes the formation of induced alkaline phosphatase protein whose rate of thermal inactivation is greatly increased (Naono and Gros, 1960). The important features of the analog effect is the fact that it occurs very quickly after administration and thereafter is reflected by a constant proportion (rather than increasing) of abnormal to normal enzyme molecules. The responsible agent, assumed to be RNA altered by incorporation of 5-fluorouracil into its structure, must be short-lived and not accumulative to account for the observed kinetics.

### B. BACTERIOPHAGE INFECTION

The original experiments (Volkin and Astrachan, 1957) with virulent phage-infected bacteria directly determined that the phage DNA-like RNA undergoes continual turnover. However, they could not distinguish whether such turnover is an obligatory part of its biologic mechanism or whether the turnover is simply a consequence of its catabolism, insofar as no ribosomal or stable RNA accumulates in this system. Further insight was gained by the discovery that some DNA-like RNA was synthesized but did not undergo turnover when protein synthesis was inhibited [chloramphenicol inhibition (Astrachan and Volkin, 1959) or deprivation of a required amino acid (Volkin, unpublished observations)]. This technique with T2 infection was more fully exploited (Volkin, 1960) by using as host the mutant B94 (Gollub and Gots, 1959) which requires both adenine and arginine. A prior incubation of T2-infected B94 cells in basal medium containing only adenine permits a definite yield of phage upon removal of the adenine and replacement with arginine alone or arginine plus deoxyadenosine (neither condition by itself will allow production of phage). During the preincubation in adenine alone, though essentially no DNA or protein is synthesized, RNA of T2 DNA-like composition steadily accumulates for about 60 to 90 minutes. Furthermore, the amount of newly formed RNA (and not any other adenine-containing cell component) can be quantitatively normalized with the subsequent phage yield. The production of T2 during the secondary incubation is limited, stopping after about 30 minutes, though the cells are capable of producing a full complement of phage if again given adenine. It seems feasible, then, to interpret these results to mean that a certain degree of stoichiometry exists between the messenger RNA and its specific product and that, as a result of its action in relaying the message, the RNA is concomitantly inactivated. Since T2 production ceases so abruptly, it must be assumed that the protein product(s) involved are not limited to the T2-induced enzymes but include phage protein itself.

## IV. ATTEMPTS AT ISOLATION OF A MESSENGER RNA

Some purification of highly radioactive RNA (presumably enriched in messenger RNA) from pulse-type experiments has been attained from both T2-infected and normally growing *E. coli* by classical centrifugation techniques (Volkin and Astrachan, 1956b, 1957) and by centrifugation through sucrose gradients (Nomura *et al.,* 1960). Since the over-all nucleotide compositions of these partially purified fractions were not unlike total RNA, it was evident that only a small degree of purification was made. In view of the observed heterogeneity of messenger RNA as it exists in the cell (Hayashi and Spiegelman, 1961), the inadequate results of such fractionations by centrifugation are not surprising. Specific hybrid formation of homologous DNA with messenger RNA as effected by Hall and Spiegelman (1961) obviously should represent a purified RNA product. On one such occasion this was proved for the T2-specific RNA banded with T2 DNA in the cesium chloride gradient (B. D. Hall and L. Astrachan, unpublished experiments). This method, by limitations in its scale, is not readily applicable to the mass isolation of messenger RNA. An ingenious technique has now been described which has been used to isolate highly purified messenger RNA and which lends itself to use on a preparative level. Bautz and Hall (1962), employing a column of denatured T4 DNA coupled to acetylated, phosphorylated cellulose, obtained specific interaction of it with only messenger RNA, the bulk of the RNA from the T4-infected cell passing through the column. The specific RNA, adsorbed to the DNA at a temperature favorable for annealing with DNA, was recovered by heating at elevated temperatures and elution. That the method is inherently specific is evident from the remarkable observation that messenger RNA specific for DNA of a T4*rII* mutant has thus been purified.

Isolation by ECTEOLA column chromatography of DNA-RNA complexes has been accomplished from *Neurospora* extracts by Schulman and Bonner (1962) and from ovarian frog egg preparations by Finamore and Volkin (1961). This technique may serve as a basis for the mass isolation of such complexes or of the specific RNA itself.

## V. *In Vitro* SYNTHESIS OF RNA

### A. POLYNUCLEOTIDE PHOSPHORYLASE

In 1955 Grunberg-Manago and Ochoa discovered an enzyme, polynucleotide phosphorylase, that for the first time allowed the cell-free, enzymatic synthesis of polyribonucleotide. Ribonucleoside diphosphates serve as substrates in the reaction and the composition of the product formed can be manipulated according to the relative concentrations of the individual diphosphates. Thus, polynucleotides can be synthesized ranging from the homopolymers to those containing all four bases in a variety of ratios (Ochoa and Heppel, 1957). It would be difficult to imagine that this mode of enzymatic synthesis is responsible for the cellular production of messenger RNA. Thus, if rigorous biologic specificity is associated with the RNA, and this specificity is a function of the nucleotide sequence, it seems unlikely that such sequences should depend on the relative pool sizes of the nucleoside diphosphate precursors. In view of the truly outstanding nature of the discovery of this enzyme, however, it is pleasing to note the eventual great usefulness of this system in the elucidation of the coding problem (see below).

### B. RNA POLYMERASE

The independent discovery from the laboratories of Weiss and Gladstone (1959), Hurwitz *et al.* (1960), and Stevens (1960) of the enzyme RNA polymerase has opened new avenues of approach not only to the question of the biologic synthesis of RNA but also to the entire question of specificity relations between RNA and protein. The enzyme has been detected in a variety of mammalian (Weiss and Gladstone, 1959; Weiss, 1962) and plant (Huang *et al.*, 1960) tissues, as well as in bacteria (Hurwitz *et al.*, 1960; Stevens, 1960; Ochoa *et al.*, 1961). RNA polymerase, like DNA polymerase (Kornberg *et al.*, 1956; Bessman *et al.*, 1958), requires DNA as a primer, and utilizes the ribonucleoside triphosphates as substrates. A net synthesis of the product can be attained. Most importantly, the composition of the DNA primer determines the composition of the RNA product. All four nucleoside triphosphates must be present for RNA synthesis, although ATP can be polymerized to polyadenylic acid (see also Spiegelman, 1959). The fact that heat-denatured DNA (Geiduschek *et al.*, 1961; Chamberlain and Berg, 1962) and single-stranded DNA of $\phi$X174 phage (Chamberlain and Berg, 1962) can serve as primers, plus the decisions based on nearest-neighbor frequency analysis (Weiss and Nakamoto, 1961), point to the conclusion that the RNA synthesized is a true complement of the primer. This enzyme, then, provides a means for the synthesis of DNA-like or messenger RNA in biologic systems.

## VI. RNA INVOLVEMENT IN *in Vitro* PROTEIN SYNTHESIS

### A. REQUIREMENT FOR DNA-DIRECTED RNA SYNTHESIS

Probably the first evidence for a synthetic pathway for RNA via nucleoside triphosphates, and, at the same time, the first indication that this RNA may be required for protein synthesis, was reported by Spiegelman (1959) with cell-free fractions from microbial systems. Spiegelman found the protoplast membrane fraction highly potent in leucine-$C^{14}$-incorporating ability into acid-insoluble protein, exceeding by a factor of 100 that found in the high-speed pellet or supernatant fraction. There was a definite requirement for all four nucleoside triphosphates, the presence of these compounds eliminating a lag in incorporation, and they could not be substituted for by the corresponding diphosphates. Nisman and Fukuhara (1959) and Nisman et al. (1961), likewise employing the *E. coli* membrane fraction, were able to assign a definite importance to the presence of intact DNA in the cell-free synthesis of β-galactosidase and the incorporation of amino acids into protein. A clear-cut approach to the question of the necessity for DNA was carried out by Novelli and his co-workers (1961). By the device of either UV-irradiation of the supernatant fraction or X-radiation of whole cells, they demonstrated an absolute requirement for the addition of DNA to their cell-free system (particles + supernatant) for the synthesis of β-galactosidase. Moreover, only DNA prepared from cells containing the gene for β-galactosidase was effective. The system required all for ribonucleoside triphosphates for enzyme synthesis and the authors invoked a mechanism for the synthesis of RNA as an intermediate, or messenger, in the transmission of information from the specific DNA to the enzyme-forming system.

Tissières *et al.* (1960) claim that the major particle site for protein synthesis in bacterial systems is the 70*S* ribosomes. Tissières has shown (Tissières and Hopkins, 1961a,b) that the RNA synthesized *in vitro* from the triphosphates has similar sedimentation properties in a sucrose gradient as that observed by Gros *et al.* (1961) following pulse labeling *in vivo*. In addition, this RNA seems to be required for the amino acid incorporation observed in the 70S ribosomes (Tissières and Hopkins, 1961b).

## B. THE USE OF POLYNUCLEOTIDES IN ESTABLISHING THE CODE

From the laboratories of Nirenberg (Matthaei and Nirenberg, 1961; Nirenberg and Matthaei, 1961; Martin *et al.,* 1962; Nirenberg *et al.,* 1962) and Ochoa (Lengyel *et al.,* 1961, 1962; Speyer *et al.,* 1962) have come the most startling applications of the microbial ribosome-supernatant amino acid incorporating systems. Nirenberg and Matthaei (1961) discovered that the addition of heat-stable RNA from different sources increased the incorporation of $C^{14}$-labeled amino acids by the ribosome-supernatant system. It is of interest that these authors note that "the correlation between the amount of incorporation and the amount of added RNA suggested stoichiometric rather than catalytic activity of the template RNA" (Nirenberg and Matthaei, 1961). In the process of testing a variety of RNA's, synthetic polynucleotides synthesized by polynucleotide phosphorylase were tested for stimulatory activity. It was found that the addition of polyuridylic acid resulted in a rather massive synthesis of poly-L-phenylalanine, and the polymerization of only this amino acid was specifically directed (Nirenberg and Matthaei, 1961). In a subsequent thorough investigation of this reaction, Nirenberg *et al.* (1962) determined that the phenylalanine was linked to soluble RNA as a true intermediate in the process. Polyuridylic acid associated in a double- or triple-stranded configuration is ineffective in amino acid stimulation, suggesting that active RNA is a single-stranded template (Nirenberg *et al.,* 1962). In any case, the observations with polyuridylic acid have been fully corroborated by Lengyel *et al.* (1961), indicating that uridylic acid residues serve as the information centers for phenylalanine insertion into protein.

The laboratories of Nirenberg and Ochoa have fully exploited the advantages of such synthetic polynucleotides as synthetic "messengers," whereby copolymers enriched in uridylic acid and containing known proportions of one or more of the other nucleotides have been tested for stimulatory activity with virtually the entire spectrum of natural amino acids. In this fashion, relatively long sequences of uridylic acid in the chain serve as templates for the production of polyphenylalanine chains, which in turn serve as "handles" (Lengyel *et al.,* 1961) for the attachment of other amino acids whose linking is directed by the specific neighboring base sequences. Thus, one may correlate the observed incorporation of a particular amino acid with the calculated frequency ratio of the uridylic sequences to the other nucleotide sequences in the polymer. Assuming a minimum coding ratio of three nucleotide sequences for the direction of a specific amino acid insertion into protein, it has been possible thus far to establish nucleotide coding letters for a larger number of amino acids. The ratio of three has a rather solid basis in terms of both the theoretical and experimental approach (see below). At the risk of publishing the score in the middle of the ball game, I have compiled the presently available "genetic code" for 17 amino acids in Table I. It should be noted that the sequential order of the nucleotides in any triplet (with the obvious exception of UUU = phenylalanine) cannot as yet be specified. In addition, the fact that more than one triplet can specify a single amino acid shows that the code is degenerate; that is, more than one coding unit can specify the insertion of a single amino acid. Thus, it should be clear that the use of enriched polyuridylic acid sequences as a "handle" probably weights the present code toward this nucleotide. A number of other triplets, absent in uridylic acid, could be equally specific for some of the amino acids listed.

TABLE I

GENETIC CODE FOR POLYMERIZATION OF 17 AMINO ACIDS[a]

| Amino acid | Triplet code letter[b] |
|---|---|
| Alanine | 1U 1C 1G |
| Arginine | 1U 1C 1G |
| Cysteine | 2U 1G |
| Glutamic acid | 1U 1C 1G |
| Glycine | 1U 2G |
| Histidine | 1U 1A 1C |
| Isoleucine | 2U 1A |
| Leucine | 2U 1C, 2U 1G |
| Lysine | 1U 2A |
| Methionine | UG[c] |
| Phenylalanine | UUU |
| Proline | 1U 2C, [CCC] |
| Serine | 2U 1C, UGC, 2U 1G |
| Threonine | 1U 2C |
| Tryptophan | 1U 2G |
| Tyrosine | 2U 1A |
| Valine | 2U 1G |

[a] Compiled from the data in Lengyel *et al.* (1961, 1962), Nirenberg and Matthaei (1961), Martin *et al.* (1962), Speyer *et al.* (1962).

[b] Sequential order unspecified.

[c] Not further specified (Martin *et al.*, 1962).

Wood and Berg (1962) have succeeded in demonstrating a significant amino acid activating effect by enzymatically synthesized RNA with a system consisting in ribosomes plus a soluble fraction that had been virtually freed of DNA. In particular, the addition of T2 phage DNA, the four ribonucleoside triphosphates and RNA polymerase resulted in a 4-fold increase in rate of reaction and a 20-fold increase in extent of reaction. Likewise, the synthesized RNA, freed of the polymerase and DNA, could activate the incorporation of amino acids some 3- to 5-fold. Although it could be shown (Chamberlain and Berg, 1962) that single-stranded DNA from heat-denatured T2 DNA, or $\phi$X174, could prime the formation of RNA by the polymerase with about equal facility as the native, or double-stranded DNA, these RNA's are completely inactive in the amino acid-incorporating system (Wood and Berg, 1962). In view of the ability of the biologically unspecific polynucleotides synthesized by polynucleotide phosphosphylase to activate amino acid incorporation in the systems of Ochoa and Nirenberg, and especially the fact that double-stranded synthetic polymers are ineffective (Nirenberg *et al.*, 1962), an enigma seems to be created on the mechanism of activation by RNA in these systems.

### C. CORRELATIONS WITH PLANT VIRUS MUTANTS

If there are suspicions about the validity of establishing a genetic code on the basis of *in vitro* incorporation of very limited amounts of radioactive amino acids into "acid-insoluble protein," it should be noted that the code letters often are the same as those surmised from *in vivo* experiments involving some plant virus mutants (see Chapter X). One nitrous acid mutant contains leucine in place of a proline residue (Tsugita and Fraenkel-Conrat, 1960). Nitrous acid acts by deaminating C to U. Inspection of Table I reveals that such a conversion of C of the proline code to U is in line with the specification of a leucine in place of proline. In the same mutant, replacement of threonine by serine can be explained in a similar way (Tsugita, 1961). In another nitrous acid mutant (No. 167) serine is replaced by phenylalanine and glutamine by valine (Tsugita, 1961). The former conversion is readily predictable from the data in Table I (2U 1C → UUU). Although no code letters have been established as yet for glutamine, it is of interest that glutamic acid (1U 1C 1G) fulfills the requirements for such a conversion to valine (2U 1G) by deamination. Also in line with the code as determined by Nirenberg and Ochoa (Table I) is the observation that the proline-phenylalanine ratio in the proteins of wild cucumber and tobacco mosaic virus can be correlated with their respective RNA C:U ratios (Yamazaki and Kaesberg, 1961).

## VII. THE TRIPLET CODE

The question of the transfer of information from nucleic acid to the synthesis of protein involves the reading of a polymer with sequences of only four different units (nucleotides) for the sequential ordering of some 20 different units (amino acids). The minimum coding unit then would be a three-letter one where 64 different sets are possible. This scheme was first proposed by Gamow (1954), who, on the basis of DNA structure, considered the code to be overlapping; that is, the letters of one triplet can be used sequentially in adjacent triplets. This coding mechanism would predict that a change in one base would result in changes in three neighboring amino acids in a protein. However, such observations that nitrous acid mutants generally are reflected in the change of only one amino acid at a time (Tsugita, 1961; Tsugita and Fraenkel-Conrat, 1960), and that the variety of abnormal human hemoglobins differ by only single amino acid changes, (Watson and Kendrew, 1961) seem to rule out the overlapping code. Helinski and Yanofsky (1962) and Henning and Yanofsky (1962) recently found that an ultraviolet-produced mutant of *E. coli* K12 contains the A protein of tryptophan synthetase differing from normal A protein by only a single amino acid substitution, glutamic acid for glycine. This same glycine is replaced by arginine in the A protein of a closely linked mutant.

Since 64 triplets can determine only 20 amino acids, it was felt that, with the exclusion of an overlapping mechanism, there must be a scheme to select the correct triplets along a continuous sequence. Crick *et al.* (1957) suggested that, although all possible sequences of amino acids can be coded, only those nucleotide triplets that sit side by side make sense, and the overlapping triplets so formed must be nonsense. In this way, they could demonstrate that the maximum number that can be coded is 20. Since then, Crick *et al.* (1961), on the basis of ingeniously conceived genetic experimentation with the B cistron of the *rII* region of bacteriophage T4, have proposed a non-overlapping triplet code where selection of the correct sequence of bases depends on reading at a fixed starting point. In this way, by starting at such a point and reading toward one direction, the insertion or deletion of a base would result in incorrect further reading of the code. However, an additional insertion or deletion (double mutant) will result in a correction of the code-reading in all sequences following the second mistake. Thus, recovery of most of the original information (in their experiments recovery to pseudo-wild type) can be obtained. It should be noted that by use of the proper triple mutants they obtained firm evidence for a probable coding ratio of three. Insofar as the region over which the suppressors they investigated covered about a quarter of the B cistron (the protein produced by this cistron may contain about 200 amino acids), the authors feel that the code is probably degenerate; otherwise, nonsense reading, which should have been encountered at a much closer distance, has been compensated for by the flexibility afforded by a degenerate code. It will be of interest to see whether as predicted (Crick *et al.*, 1961), the protein produced by some of the double mutants has been altered in amino acids through the length of the polypeptide chain corresponding to the region between the two alterations in the DNA chain.

## VIII. SUMMARY

An RNA of composition similar to DNA has been detected in a number of different biologic systems. Its presence in phage-infected bacteria is particularly apparent, since ribosomal RNA synthesis is prevented in that system. This RNA probably has a sedimentation constant of 12–14S but appears to be largely reversibly associated with ribosomal particles. The RNA undergoes rapid turnover, which can be measured by its molecular degradation in the phage-infected cell, and by its translocation to other subcellular components in actively growing bacteria. The turnover phenomenon seems to be part and parcel of its function as deduced from kinetic studies. Moreover, by correlating the amount of such RNA synthesis with (*1*) phage production, (*2*) enzyme synthesis, and (*3*) *in vitro* amino acid incorporation, a definite stoichiometry has been found to exist between the amount synthesized and the protein product. Specific RNA-DNA hybrids (where the RNA has a composition similar to the DNA) can be isolated from extracts, as well as produced by "annealing" of RNA with only the homologous DNA. RNA of DNA-like composition likewise can be synthesized *in vitro* with the enzyme, RNA polymerase. RNA synthesized in this manner markedly activates the incorporation of amino acids into protein in cell-free systems. In all probability such *in vitro*-synthesized, complete, RNA molecules will be shown eventually to be capable of directing the synthesis of complete, biologically active proteins. Presently, however, it is possible to use synthetic polynucleotides to demonstrate the existence of a definite specificity between the nucleotide composition and the particular amino acid incorporated into protein. In this way, a partial nucleotide triple code for most of the naturally occurring amino acids has been resolved.

## REFERENCES

Astrachan, L. (1960). *Federation Proc.* 19, 198.
Astrachan, L., and Fisher, T. N. (1961). *Federation Proc.* 20, 359.
Astrachan, L., and Volkin, E. (1958). *Biochim. et Biophys. Acta* 29, 536.
Astrachan, L., and Volkin, E. (1959). *Biochim. et Biophys. Acta* 32, 449.
Bautz, E. K. F., and Hall, B. D. (1962). *6th Annual Meeting Biophys. Soc., Washington, D.C.* Abstr. TA8.
Bessman, M. J. (1959). *J. Biol. Chem.* 234, 2735.

Bessman, M. J., Lehman, I. R., Simms, E. S., and Kornberg, A. (1958). *J. Biol. Chem.* **233,** 171.

Boezi, J. A. Bolton, E. T., Britten, R. J., Cowie, D. B., McCarthy, B. J., Midgley, J. E., and Roberts, R. B. (1961). *Carnegie Inst. Wash. Yearbook* **60,** 263.

Boezi, J. A., and Cowie, D. B. (1961). *Biophys. J.* **1,** 639.

Boulanger, P., and Montreuil, J. (1952). *Biochim. et Biophys. Acta* **9,** 619.

Brachet, J. (1933). *Arch. biol (Liège)* **44,** 519.

Brachet, J. (1937). *Arch. biol. (Liège)* **48,** 529.

Brachet, J. (1940a). *Arch. biol. (Liège)* **51,** 151, 167.

Brachet, J. (1940b). *Compt. rend. soc. biol.* **133,** 88, 90.

Brenner, S., Jacob, F., and Meselson, M. (1961). *Nature* **190,** 576.

Bussard, A., Naono, S., Gros, F., and Monod, J. (1960). *Compt. rend. acad. sci.* **250,** 4049.

Caspersson, T. (1941). *Naturwiss.* **29,** 33.

Chamberlain, M., and Berg, P. (1962). *Proc. Natl. Acad. Sci. U.S.* **48,** 81.

Cohen, S. S., Barnes, H. B., and Lichtenstein, J. (1961). *J. Biol. Chem.* **236,** 1448.

Cohn, W. E. (1949). *Science* **109,** 377.

Cohn, W. E. (1950a). *J. Am. Chem. Soc.* **72,** 1471.

Cohn, W. E. (1950b). *J. Am. Chem. Soc.* **72,** 2811.

Cohn, W. E. (1951). *J. Am. Chem. Soc.* **73,** 1539.

Crick, F. H. C., Griffith, J. S., and Orgel, L. E. (1957). *Proc. Natl. Acad. Sci. U.S.* **43,** 416.

Crick, F. H. C., Barnett, L., Brenner, S., and Watts-Robin, R. J. (1961). *Nature* **192,** 1227.

Finamore, F. J., and Volkin, E. (1961). *Am. Chem. Soc. Div. of Biol. Chem. Abstr.* 53C.

Flaks, J. G., Lichtenstein, J., and Cohen, S. S. (1959). *J. Biol. Chem.* **234,** 1507.

Gamow, G. (1954). *Nature* **173,** 318.

Geiduscheck, E. P., Nakamoto, T., and Weiss, S. B. (1961). *Proc. Natl. Acad. Sci. U.S.* **47,** 1405.

Gollub, E. G., and Gots, J. S. (1959). *J. Bacteriol.* **78,** 320.

Gros, F., Hiatt, H., Gilbert, W., Kurland, C. G., Riseborough, R. W., and Watson, J. D. (1961). *Nature* **190,** 581.

Grunberg-Manago, M., and Ochoa, S. (1955). *J. Am. Chem. Soc.* **77,** 3165.

Hall, B. D., and Spiegelman, S. (1961). *Proc. Natl. Acad. Sci. U.S.* **47,** 137.

Hayashi, M., and Spiegelman, S. (1961). *Proc. Natl. Acad. Sci. U.S.* **47,** 1564.

Helinski, D. R., and Yanofsky, C. (1962). *Proc. Natl. Acad. Sci. U.S.* **48,** 173.

Henning, U., and Yanofsky, C. (1962). *Proc. Natl. Acad. Sci. U.S.* **48,** 183.

Hokin, L. E., and Hokin, M. R. (1954). *Biochim. et Biophys. Acta* **13,** 401.

Huang, R. C., Maheswari, N., and Bonner, J. (1960). *Biochem. Biophys. Research Communs.* **3,** 689.

Hurwitz, J., Bresler, A., and Biringer, R. (1960). *Biochem. Biophys. Research Communs.* **3,** 15.

Jacob, F., and Monod, J. (1961). *J. Mol. Biol.* **3,** 318.

Keck, K., Mahler, H. R., and Fraser, D. (1960). *Arch. Biochem. Biophys.* **86,** 85.

Kitazume, Y., Ycas, M., and Vincent, W. S. (1962). *Proc. Natl. Acad. Sci. U.S.* **48,** 265.

Kornberg, A., Lehman, I. R., and Simms, E. S. (1956). *Federation Proc.* **15,** 291.

Kornberg, A., Zimmerman, S. B., Kornberg, S. R., and Josse, J. (1959). *Proc. Natl. Acad. Sci. U.S.* **45,** 772.

Lengyel, P., Speyer, J. F., and Ochoa, S. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 1936.

Lengyel, P., Speyer, J. F., Basilio, C., and Ochoa, S. (1962). *Proc. Natl. Acad. Sci. U.S.* **48**, 282.

Logan, R., and Davidson, J. N. (1957). *Biochim. et Biophys. Acta* **24**, 196.

Martin, R. G., Matthaei, J. H., Jones, O. W., and Nirenberg, M. W. (1962). *Biochem. Biophys. Research Communs.* **6**, 410.

Matthaei, J. H., and Nirenberg, M. W. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 1580.

Miescher, F. (1871). *Hoppe-Seyler's Med. Chem. Unters.* p. 441.

Monod, J. (1958). *Rec. trav. chim.* **77**, 569.

Naono, S., and Gros, F. (1960). *Compt. rend. acad. sci.* **250**, 3527, 3889.

Nirenberg, M. W., and Matthaei, J. H. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 1588.

Nirenberg, M. W., Matthaei, J. H., and Jones, O. W. (1962). *Proc. Natl. Acad. Sci. U.S.* **48**, 104.

Nisman, B., and Fukuhara, H. (1959). *Compt. rend. acad. sci.* **250**, 410.

Nisman, B., Kayser, A., Demailly, J., Jr., and Genin, C. (1961). *Compt. rend. acad. sci.* **252**, 2619.

Nomura, M., Hall, B. D., and Spiegelman, S. (1960). *J. Mol. Biol.* **2**, 306.

Novelli, G. D., Kameyama, T., and Eisenstadt, J. M. (1961). *J. Cellular Comp. Physiol.* Suppl. 1, **58**, 225.

Ochoa, S., and Heppel, L. A. (1957). *In* "The Chemical Basis Heredity" (W. D. McElroy and B. Glass, eds.), p. 615. Johns Hopkins Press, Baltimore.

Ochoa, S., Burma, D. P., Kroger, H., and Weill, J. D. (1961). *Federation Proc.* **20**, 362.

Pardee, A. B., Jacob, F., and Monod, J. (1959). *J. Mol. Biol.* **1**, 165.

Riley, M., Pardee, A. B., and Monod, J. (1960). *J. Mol. Biol.* **2**, 216.

Roberts, R. B., Britten, R. J., and McClure, F. T. (1961). *Biophys. J.* **1**, 649.

Schulman, H. M., and Bonner, D. M. (1962). *Proc. Natl. Acad. Sci. U.S.* **48**, 53.

Sommerville, R., and Greenberg, G. R. (1959). *Federation Proc.* **18**, 327.

Speyer, J. F., Lengyel, P., Basilio, C., and Ochoa, S. (1962). *Proc. Natl. Acad. Sci. U.S.* **48**, 63.

Spiegelman, S. (1959). *In* "Recent Progress in Microbiology," p. 81. *7th Intern. Congr. for Microbiol., Stockholm, 1958.* Almquist and Wiksell, Uppsala.

Spiegelman, S., Hall, B. D., and Storck, R. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 1135.

Stevens, A. (1960). *Biochem. Biophys. Research Communs.* **3**, 92.

Tissières, A., and Hopkins, J. W. (1961a). *Cold Spring Harbor Symposia Quant. Biol.* **26**, 126.

Tissières, A., and Hopkins, J. W. (1961b). *Proc. Natl. Acad. Sci. U.S.* **47**, 2015.

Tissières, A., Schlessinger, D., and Gros, F. (1960). *Proc. Natl. Acad. Sci. U.S.* **46**, 1450.

Tsugita, A. (1961). *Protein, Nucleic Acid Enzyme (Tokyo)* **6**, 385.

Tsugita, A., and Fraenkel-Conrat, H. (1960). *Proc. Natl. Acad. Sci. U.S.* **46**, 636.

Vischer, E., and Chargaff, E. (1947). *J. Biol. Chem.* **168**, 781.

Vischer, E., and Chargaff, E. (1948a). *J. Biol. Chem.* **176**, 703, 715.

Vischer, E., and Chargaff, E. (1948b). *Federation Proc.* **7**, 197.

Volkin, E. (1959). *In* "Biochemistry of Viruses" (E. Broda and W. Frisch-Niggemeyer, eds.), pp. 7, 212. *Proc. 4th Intern. Congr. Biochem., Vienna, 1958.* Pergamon Press, New York.

Volkin, E. (1960). *Proc. Natl. Acad. Sci. U.S.* **46**, 1336.

Volkin, E. (1962). *Federation Proc.* **21**, 112.

Volkin, E., and Astrachan, L. (1956a). *Virology* **2**, 149.
Volkin, E., and Astrachan, L. (1956b). *Virology* **2**, 433.
Volkin, E., and Astrachan, L. (1957). *In* "The Chemical Basis of Heredity," (W. D. McElroy and B. Glass, eds.), p. 686. Johns Hopkins Press, Baltimore.
Volkin, E., Astrachan, L., and Countryman, J. L. (1958). *Virology* **6**, 545.
Watson, H. C., and Kendrew, J. C. (1961). *Nature* **190**, 670.
Weiss, S. B. (1962). *Federation Proc.* **21**, 120.
Weiss, S. B., and Gladstone, L. (1959). *J. Am. Chem. Soc.* **81**, 4118.
Weiss, S. B., and Nakamoto, T. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 1400.
Wood, W. B., and Berg, P. (1962). *Proc. Natl. Acad. Sci. U.S.* **48**, 94.
Yamazaki, H., and Kaesberg, P. (1961). *Nature* **191**, 97.
Yčas, M., and Vincent, W. S. (1960). *Proc. Natl. Acad. Sci. U.S.* **46**, 804.
Zamecnik, P. C. (1960). *Harvey Lectures* **54**, 256.

## *Chapter* VII

# Kinetic Studies of the Synthesis of RNA and Ribosomes

R. B. ROBERTS, R. J. BRITTEN, AND
B. J. McCARTHY

*Carnegie Institution of Washington,*
*Department of Terrestrial Magnetism,*
*Washington, D. C.*

I. Introduction . . . 292
  A. Microsomes and Ribosomes . . . 292
  B. Effects of Magnesium . . . 293
  C. Other Influences on Ribosome Patterns . . . 296
  D. Generality of Ribosome Patterns . . . 298
  E. Ribosomes as Site of Protein Synthesis . . . 299
II. Properties of Ribosomes *in Vitro* . . . 299
  A. Preparation and Purification of Ribosomes . . . 299
  B. Protein/Nucleic Acid Ratio . . . 300
  C. Protein Components . . . 302
  D. Nucleic Acid Components . . . 305
  E. Structure of Ribosomes . . . 306
III. Fractionation of Ribosomes . . . 308
  A. Chromatography . . . 308
  B. Sedimentation Analysis . . . 309
  C. Density Gradient Banding . . . 310
  D. Electrophoresis . . . 310
  E. Two-Phase Systems . . . 310
  F. Conclusion . . . 310
IV. Metabolic Pools . . . 311
V. Kinetics of Incorporation into RNA of Ribosomes . . . 315
  A. DEAE Chromatography . . . 315
  B. Sedimentation Analysis . . . 316
  C. Quantitative Kinetic Analysis . . . 321
  D. Stages of Synthesis of the Ribosomal RNA . . . 326
VI. Kinetic Studies of Phenol-Extracted RNA . . . 329
VII. Kinetics of Incorporation into Protein . . . 331
  A. Incorporation of $C^{14}$-Leucine into the Ribosomes . . . 331
  B. Effect of Chloramphenicol on Ribosome Synthesis . . . 333
  C. Assembly of the Ribosomes . . . 333
VIII. Composition of Newly Formed RNA . . . 335
  A. Ambiguities . . . 335
  B. Time Course of Composition . . . 337
  C. Fractionation . . . 338

D. RNA Classes . . . . . . . . . . . . . . . . . 340
E. Mechanisms for the Change in Apparent Composition . . . . . 341
IX. Discussion . . . . . . . . . . . . . . . . . 342
A. Role of Ribosomes in Information Transfer . . . . . . . 342
B. Correlation with Enzyme Induction . . . . . . . . . 343
C. Messenger Theory . . . . . . . . . . . . . 344
D. Homogeneity of Ribosomes . . . . . . . . . . . 345
E. Site of Ribosome Synthesis . . . . . . . . . . . 346
F. DNA-RNA Hybrids . . . . . . . . . . . . . 347
G. Self-Replication of Ribosomes . . . . . . . . . . 347
H. Cell-Free Synthesis of Ribosomes . . . . . . . . . 348
I. State of Ribosomes in Living Cells . . . . . . . . . 348
X. Epilogue . . . . . . . . . . . . . . . . . 350
References . . . . . . . . . . . . . . . . . 350

## I. INTRODUCTION

### A. MICROSOMES AND RIBOSOMES

The term "microsome" was originated by Claude in 1943 to describe a fraction obtained by differential centrifugation of disrupted vertebrate cells (Claude, 1943). After breakage, the cells were centrifuged at low *g* to sediment nuclei, cell walls, and mitochondria. The supernatant fluid, devoid of structures visible in the microscope, was then centrifuged at 100,000 *g* for 2 hours and the pellet so obtained was designated the microsomal fraction. In it appeared much of the cytoplasmic ribonucleic acid (RNA) and lipid.

Electron microscopy of this pellet showed that it consisted mainly of membranes to which were attached dense particles of about 200 Å diameter. Electron micrographs of the same cells showed an extensive network of membranes designated the endoplasmic reticulum. In addition, dense particles could be seen sometimes free and sometimes attached to the membranes. The RNA content of the pellet was attributed to its content of particles (see review, Palade, 1958). Particles from bacteria had been observed earlier (Luria *et al.*, 1943.)

Analysis of the microsome fraction in the analytical centrifuge, especially after treatment with deoxycholate to dissolve the membranous material, showed the presence of objects having sedimentation coefficients in the range 20–100S (Peterman and Hamilton, 1957).

Similar fractions of high RNA content were obtained from bacteria, yeast, and plant cells. All showed the presence of a number of discrete peaks in the analytical centrifuge having sedimentation coefficients of 20–100S (Schachman *et al.*, 1952; Chao and Schachman, 1956; Ts'O *et al.*, 1956).

Since the microsomal fraction included both the membrane material and the RNA-rich particles as well as any protein which was sedimented by the prolonged centrifugation, the reported composition of the microsomes was highly variable. In 1957 the term "ribosome" (originally suggested by H. M. Dintzis) was introduced to distinguish the particulate material from the remainder of the microsome fraction (Roberts, 1958).

### B. EFFECTS OF MAGNESIUM

In the same year the effect of magnesium concentration upon the sedimentation coefficients obtained for the ribosomes was recognized. Chao showed that the larger ribosomes (about 80S) obtained from yeast

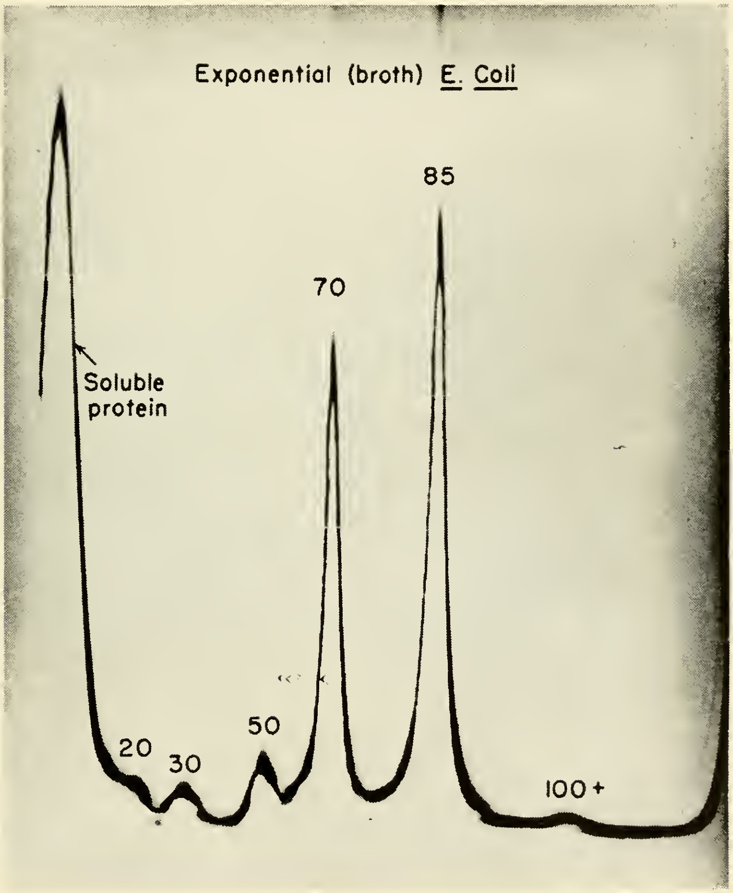


FIG. 1. Centrifugation pattern of juice from *E. coli* growing in broth medium. These rapidly growing cells have a high content of ribosomes. Sedimentation from left to right observed in Spinco Model E ultracentrifuge.

dissociated into two components of about 60S and 40S when the magnesium concentration of the suspending fluid was reduced from $10^{-3}$ $M$ (Chao, 1957). Bolton *et al.* independently found a similar effect in the ribosomes of *Escherichia coli* (Bolton *et al.*, 1958). Ts'O *et al.* reported the dissociation of the 80S component of ribosomes of pea seedlings into components of 60S, 40S, and 26S (Ts'O *et al.*, 1958). A Mg/P ratio of 0.3 was required to preserve the integrity of 80S particles from pea seedlings (Edelman *et al.*, 1960).

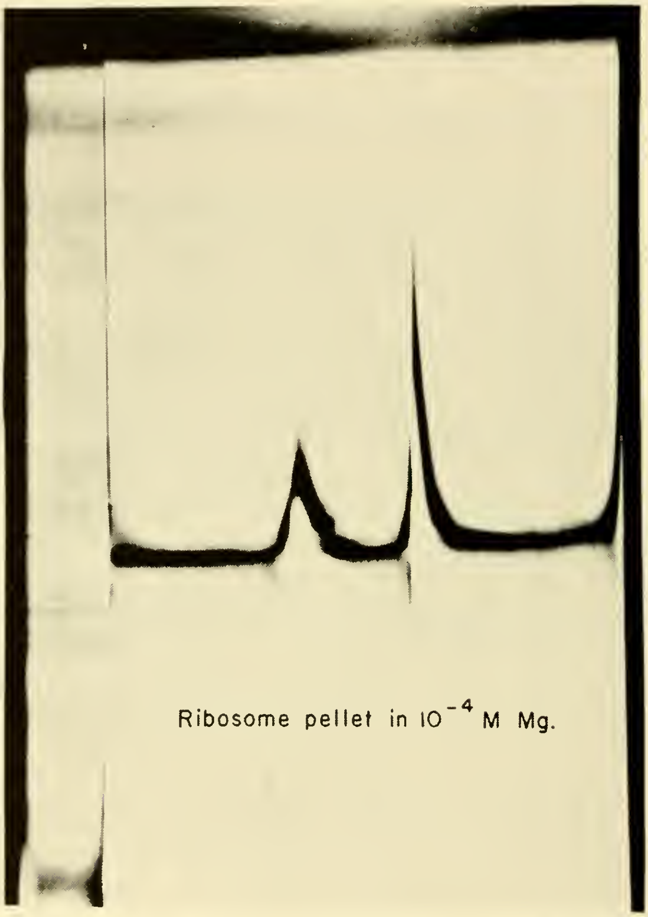


FIG. 2. Centrifugation pattern of ribosome pellet in $10^{-4}$ $M$ Mg. Only 50S and 30S ribosomes appear.

The sedimentation coefficients reported in early papers vary widely; most of the values were not corrected for viscosity or concentration.

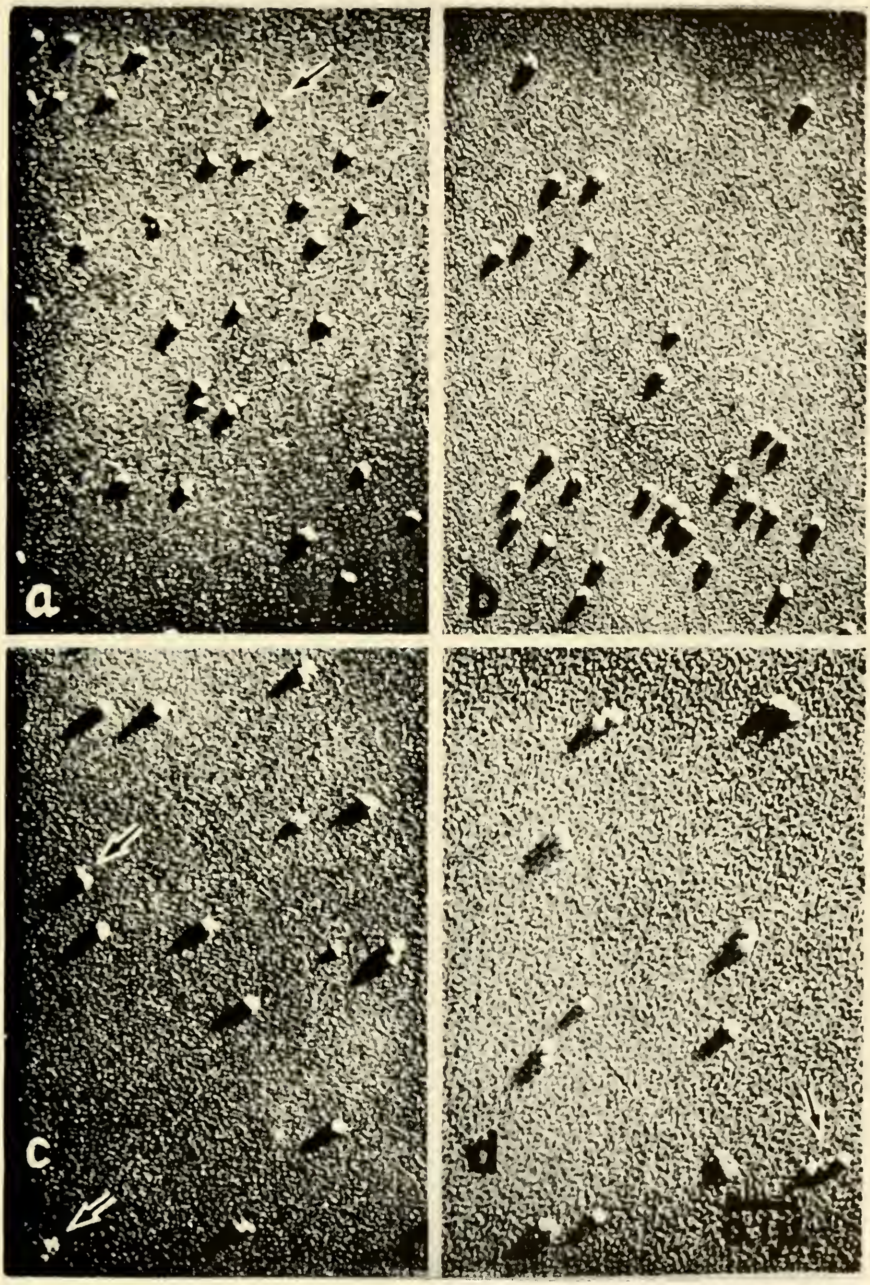


FIG. 3. Ribosomes of *E. coli* (×100,000). (Hall and Slayter, 1959.) (*a*) 30S preparation, (*b*) 50S preparation, (*c*) 70S preparation, (*d*) 100S preparation.

When these corrections are made the ribosomes of *E. coli* show three prominent groups at 30.6 ± 1S, 50.0 ± 1S, and 69.1 ± 1S. These are usually rounded off to 30, 50, and 70S for identification (Tissières *et al.*, 1959).

Ribosomes extracted from *E. coli* at $10^{-2}$ *M* Mg show predominantly the 70S (or larger) form (Fig. 1), but when the extraction is carried out in $10^{-4}$ Mg or when the Mg concentration is reduced to $10^{-4}$ *M* after extraction the large particles give rise to equal numbers of 30S and 50S (Fig. 2). This fact, together with electron microscope pictures which show shapes consistent with a 30–50S complex (Fig. 3), (Hall and Slayter, 1959; Huxley and Zubay, 1960) indicates that the 30S and 50S ribosomes associate reversibly to form 70S particles.

In addition, ribosomes of sedimentation coefficients greater than 70S are observed. One of these (about 100S) appears to be a dimer of the 70S particle (Fig. 3). The reactions are not simple, however. An intermediate group (called 85S for identification) also appears and its sedimentation coefficient as well as its proportions depends on the magnesium concentration (Britten and McCarthy, 1959). As this 85S group yields 30S and 50S particles in equal numbers when the magnesium concentration is reduced, it appears to be due either to an altered shape or hydration of the 70S or 100S particles or to a rapid equilibration between the 70S and 100S forms which yields an intermediate sedimentation coefficient.

### C. OTHER INFLUENCES ON RIBOSOME PATTERNS

The sedimentation pattern also depends upon the metabolic state of the cell at the time of breaking and upon the method of breaking the cells (Fig. 4). McCarthy has shown that cells rapidly accumulate the 100S particles when their energy supply (glucose) is exhausted and revert with equal rapidity to the normal pattern when glucose is restored (McCarthy, 1960). Similar changes also appear when mutants are deprived of a required metabolite. There is little difference, however, in ribosomes taken from different phases of the growth cycle in synchronized cultures (Britten *et al.*, 1960).

Prolonged incubation of growing cells in Mg-free media reduces the ribosome content to less than 5%. The ribosome content is restored (following an exponential growth curve) when the magnesium of the medium is replenished (McCarthy, 1962).

The ribosome pattern does not result simply from an equilibrium among two basic units and combination thereof. Some 30S and 50S particles are present in the cell juice even when the magnesium concentration is high. These are sometimes designated "native" 30S and 50S ribosomes in distinction to the "derived" 30S and 50S particles which are obtained from the dissociation of 70S particles. Native particles have higher specific radioactivities (after a short period of exposure to the tracer) than do the derived particles (Britten and McCarthy, 1959). Isolated native particles do not combine (Green and Hall, 1961).

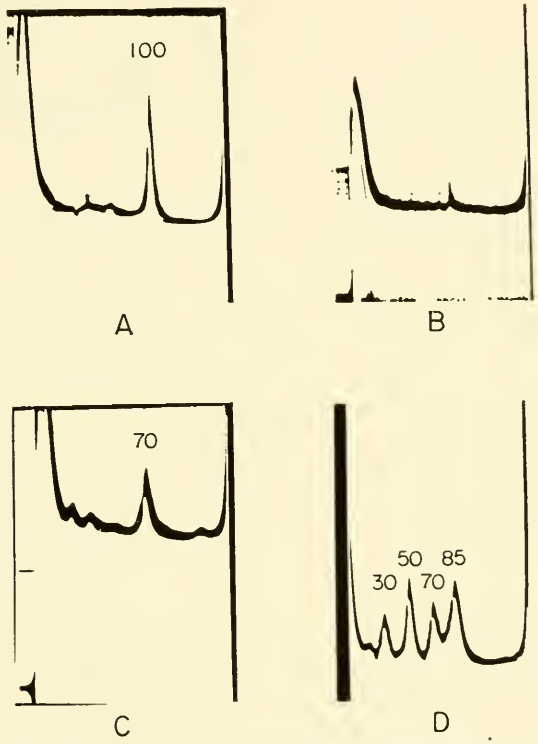


FIG. 4. Changes in ribosome pattern caused by different metabolic conditions: A. The 100S component is prominent when growth is stopped by lack of energy source. B. Magnesium starvation causes disappearance of ribosomes. C. Incubation with chloramphenicol results in accumulation of small particles. Ribosomes already present appear as 70S. D. Incubation with 5-fluorouracil results in accumulation of 30S and 50S ribosomes.

Furthermore, it has been shown *in vivo* (McQuillen *et al.*, 1959) and *in vitro* (Tissières *et al.*, 1960) that the 70S particles carrying newly made protein do not dissociate completely when the magnesium concentration is reduced to $10^{-4}$ $M$.

In addition to the main peaks which are readily observed smaller groups of particles sometimes appear. Many preparations show a peak at about 20S. The 20S component is complex, containing one component rich in protein and another rich in RNA which may be a ribosome precursor. The 20S particle also appeared after prolonged storage of ribosomes at $10^{-4}$ $M$ magnesium. In this circumstance a sufficient quantity was obtained to show that the protein/nucleic acid ratio was about 40/60 but it is not known whether this degradation product is the same as the particle observed in cell juices (Roberts and Duerksen, 1960; Aronson and McCarthy, 1961).

Another particle about 43S plays a prominent role as a protein deficient precursor to the 50S ribosome (see Section V). It has also been observed in the juice of cells in which DNA synthesis was prevented by lack of thymine (Roberts, 1960) and as a degradation product of 50S ribosomes (Elson, 1961).

Abnormal patterns develop when cells are exposed to antibiotics or analogs. Chloramphenicol causes the accumulation of small particles deficient in protein while preventing the entry of $P^{32}$ into the normal particles (Pardee *et al.*, 1957; Nomura and Watson, 1959; Bolton, 1959). 5-Fluorouracil causes the accumulation of 30S and 50S ribosomes which are unable to form 70S particles (Aronson, 1961).

A part of the complexity of the ribosome pattern may be due to the association of newly synthesized RNA or protein with the particles (see Section V,C and VIII). Such material might facilitate or inhibit association of the basic units. In most experiments its presence would be undetected and uncontrolled, thus giving rise to unpredictable variations.

### D. GENERALITY OF RIBOSOME PATTERNS

Because of the wide range of ribosome patterns which can be observed in one organism it is difficult to compare the ribosomes from different sources. Those which have been most extensively studied show in common particle groups of 70–100S which dissociate in low concentrations of magnesium into two groups of 26–40S and 40–60S. McCarthy has surveyed the ribosomes derived from a number of microorganisms and found a general similarity of patterns (McCarthy, 1959). Some minor differences appeared but they cannot be considered significant without extensive studies of the effects of growth conditions. Also, the magnesium concentration required for stability varies from one organism to another.

In view of these difficulties of intercomparing ribosomes and because most of the studies of ribosome synthesis have been carried out with *E. coli*, it seems best to describe the process of ribosome synthesis as it has been observed in this organism. We hope that this will prove to be a specific example of a general process. Furthermore, we have omitted any reference to the synthesis of RNA in virus-infected cells as the relation of this process to ribosome synthesis is still obscure.

### E. RIBOSOMES AS SITE OF PROTEIN SYNTHESIS

Interest in the biosynthesis of ribosomes stems primarily from the belief that the ribosomes serve as the principal machinery of the cell for protein synthesis. The evidence favoring this view has accumulated over a period of years and is now quite convincing. Cassperson and Brachet observed that RNA was invariably present in high quantity wherever or whenever there was a high rate of synthesis (Cassperson, 1950; Brachet, 1950). Caldwell and Hinshelwood showed a quantitative correlation between the quantity of RNA and the growth rate of a wide range of different cells and different conditions (Caldwell and Hinshelwood, 1950). Subsequently, when it was found that most of the RNA was located in ribosomes, the same evidence indicated that a high concentration of ribosomes was correlated with a high rate of protein synthesis.

Further evidence appeared in the studies of incorporation of labeled amino acids into protein. *In vivo* experiments showed that the radioactivity of the microsome fraction has a rapid initial rise, whereas that of the soluble proteins has an initial delay suggesting a precursor product relationship (Borsook *et al.*, 1950; Littlefield *et al.*, 1955). In mammalian tissues the times involved were about 15 minutes and in bacteria the times were about 15 seconds. Furthermore, with the bacteria it was possible to follow the "pulse" of radioactivity with a "chase" of nonradioactive material and to show a transfer of radioactivity from the ribosomes to the soluble protein (McQuillen *et al.*, 1959).

*In vitro* experiments showed that ribosomes were an essential component of any cell-free system capable of incorporating amino acids into protein (Littlefield and Keller, 1957; Tissières *et al.*, 1960). In these cell-free preparations a large part of the radioactivity incorporated into peptide linkage remained associated with the ribosomes.

It is now widely accepted that ribosomes furnish the sites of protein synthesis and that the information for the assembly of the amino acids in their proper order is transferred either to the RNA of the ribosomes or to an RNA which is associated with the ribosomes while it acts as template for protein synthesis.

## II. PROPERTIES OF RIBOSOMES *in Vitro*

### A. PREPARATION AND PURIFICATION OF RIBOSOMES

The preparation of ribosomes begins with the harvesting of the cells. A culture in the desired metabolic condition is rapidly chilled to maintain that state and centrifuged. The resulting pellet is washed three times to replace the growth medium by a buffer suitable for subsequent processing of ribosomes. The buffer most frequently used is tris(hydroxymethyl)aminomethane 0.01 $M$ pH 7.4 with magnesium added as the chloride at the desired concentration (usually $10^{-2}\,M$).

The cells may then be broken by grinding the pellet with twice its weight of alumina, by repeated freezing and thawing with lysozyme, or by forcing them through a small orifice in the French pressure cell. The pressure cell is preferable for the preparation of ribosomes as it is highly efficient in breaking most cells and the DNA is degraded so that the cell juice is not viscous.

The cell juice is then diluted to less than 100 mg wet weight/ml and the cell walls, membranes, and any unbroken cells are removed by centrifugation for 1 minute at 40,000 rpm in the angle head rotor of the Spinco Model L ultracentrifuge. The supernatant fluid (40K 1′S) is decanted and the pellet is either discarded or resuspended to recover its content of ribosomes (about 10–20%).

The crude ribosome pellet is then obtained by centrifugation of the 40K 1′S. If complete recovery of ribosomes as small as 30S is required, 2–3 hours of centrifugation is needed. Alternatively, if only a 90% recovery of 70–100S ribosomes is desired, 30 minutes' centrifugation will suffice.

The first ribosome pellet is highly contaminated, partly by small bits of cell wall and membrane and partly by soluble protein and RNA. Much of the contaminating material will not resuspend readily and can be removed by a brief centrifugation (40K 1′). Alternate short and long cycles of centrifugation helps to reduce the contamination and the ribosome pellet becomes more colorless and transparent. By appropriate choice of the times of centrifugation, pellets rich in one or another class of ribosome can be obtained. However, this technique is not adequate for measurements of the protein or enzyme content of ribosomes. Further purification of the pellet by sedimentation through sucrose gradients or by column chromatography shows that after three cycles of sedimentation in the angle head rotor as much as half the protein content of the pellet may be due to contamination. The best use of differential centrifugation is to prepare partially purified ribosomes as the starting material for other techniques of separation.

### B. PROTEIN/NUCLEIC ACID RATIO

Various values reported for the protein and nucleic acid content of ribosomes are listed in Table I. As many of these experiments were carried out before the difficulties of removing protein contaminants were fully appreciated, the protein content is likely to be overestimated. There is no certainty that the protein/nucleic acid ratio varies in

TABLE I
RNA CONTENT OF RIBOSOMES

| Source of ribosomes | RNA (%) | Reference |
|---|---|---|
| Calf liver | 40 | Peterman and Hamilton (1957) |
| Yeast | 42 | Chao (1957) |
| Pea seedlings | 40 | Ts'O *et al.* (1956) |
| Rabbit reticulocytes | 50 | Dintzis *et al.* (1958) |
| *Escherichia coli* | 60 | Roberts *et al.* (1958) |
| *Escherichia coli* | 63 | Tissières *et al.* (1959) |

different species; the reported variations may be due to variable contamination.

Ribosomes of different sedimentation constants appear to have the same protein/nucleic acid ratio. Figure 5 shows the correspondence between the protein (measured by $S^{35}$) and the nucleic acid (measured by

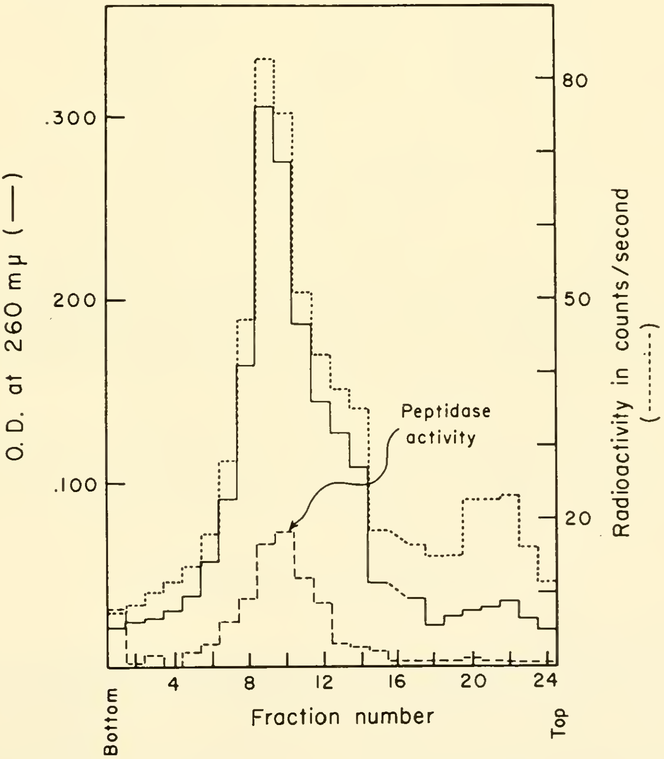


FIG. 5. Sedimentation analysis of a ribosome pellet shows correspondence between nucleic acid content indicated by optical density (O.D.), protein indicated by $S^{35}$, and peptidase activity.

optical density) in ribosomes separated by sedimentation. Similar results are shown in Fig. 6 using $P^{32}$ and $C^{14}$-leucine. Thus the ribosomes of *E. coli*, whether 30S, 50S, 70S, or 100S, all have the same nucleic acid content of 63% or possibly higher. Notable exceptions to this statement are the particles which accumulate during growth in chloramphenicol and the precursors of the ribosomes (see Sections V–VII).

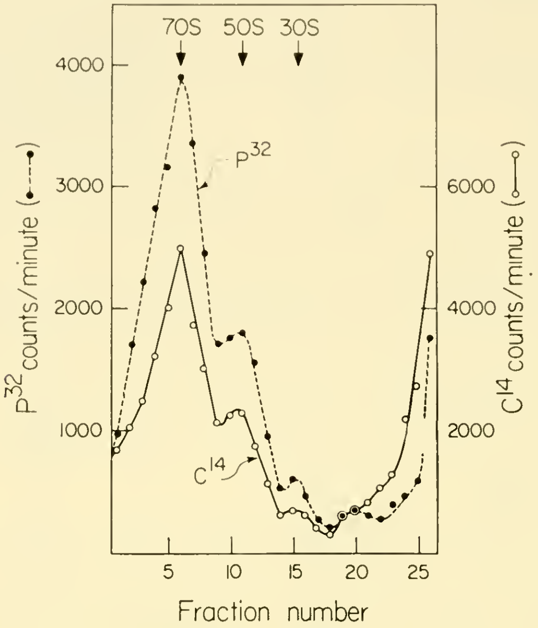


FIG. 6. Sedimentation analysis of a total cell extract from *E. coli* cells in tris-HCl 0.01 *M*, pH 7.4, $MgCl_2$ 0.01 *M*, randomly labeled with $C^{14}$-leucine and $P^{32}O_4^{---}$. Centrifugation 90 minutes at 37,000 rpm, 4°C. Note same ratio of protein/nucleic acid in ribosomes of different sedimentation coefficients. Protein indicated by radioactivity of $C^{14}$-leucine, nucleic acid indicated by $P^{32}$.

## C. PROTEIN COMPONENTS

### 1. *Structural Elements*

The protein and nucleic acid components of ribosomes are readily dissociated either by treatment with 4 *M* urea or by removal of magnesium with chelating agents such as disodium ethylenediaminetetraacetate. Bolton showed the presence of several components separable by chromatography (Bolton, 1958). More extensive studies by Waller and Harris show at least 20 components which are resolved by electrophoresis.

Methionine (about 47%), alanine (about 37%), and serine (about 11%) account for most of the $NH_2$ terminal amino acids and the average molecular weight of the individual protein components estimated from the proportion of $NH_2$ terminal amino acids is 25,000 (Waller and Harris, 1961). Since the molecular weight of the total protein of the 30S particle is only 260,000, any one 30S particle could include only about 10 of the observed components.

### 2. *Ribosomal Enzymes*

Latent ribonuclease (RNase) was found in chemically isolated ribonucleoprotein of *E. coli*. The enzyme was not active until the structure of the nucleoprotein was disrupted (Elson, 1958, 1959).

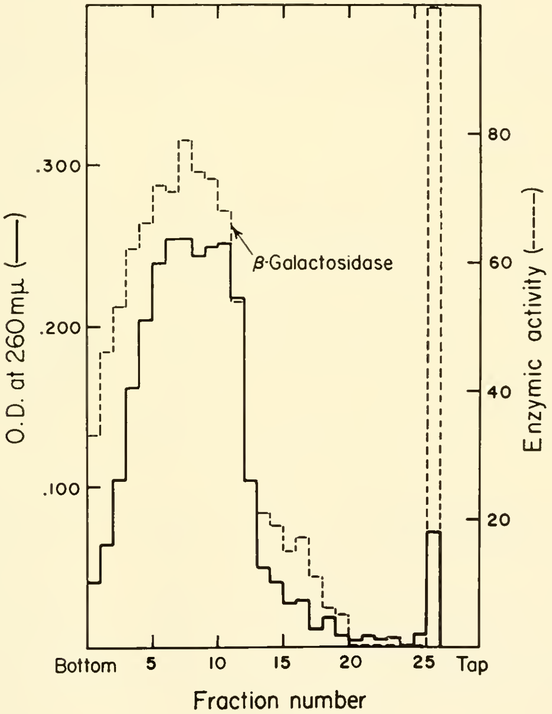


FIG. 7. Sedimentation analysis shows a correspondence between ribosomes indicated by optical density (O.D.) and β-galactosidase activity. (Cowie *et al.*, 1961.)

Latent RNase also appeared in ribosomes isolated by centrifugation or chromatography. Furthermore, the total cellular content of RNase could be attributed to the RNase of the ribosomes (Bolton, 1958). The enzyme is either localized in 30S particles exclusively (Elson and Tal,

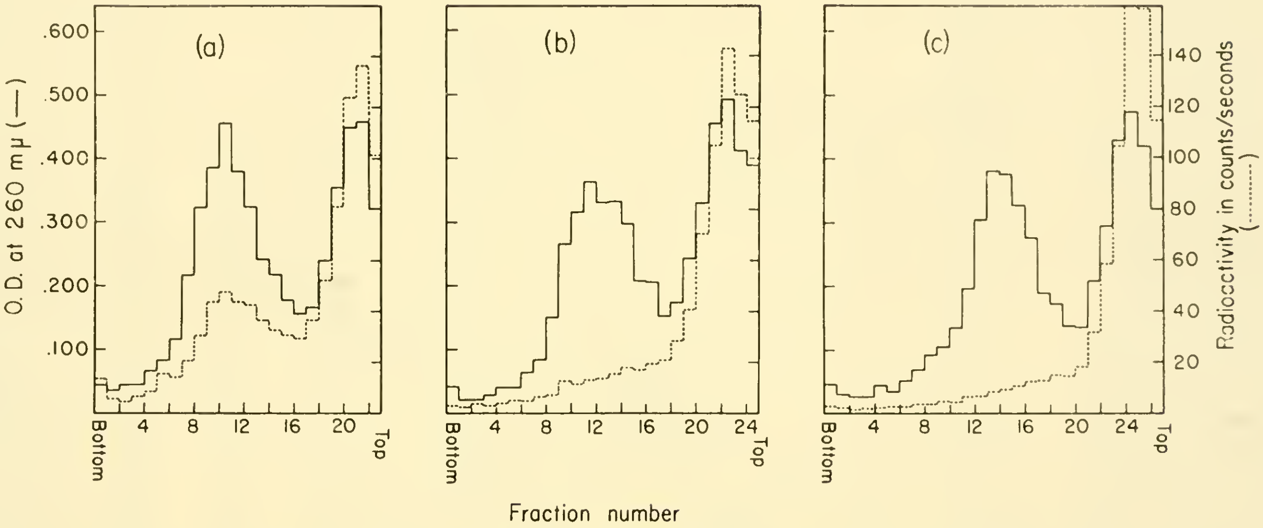


FIG. 8. Transient association of newly incorporated $S^{35}$ with ribosomes shown by sedimentation analysis of total cell juice. (*a*) Cells incubated 15 seconds with $S^{35}O_4^{--}$, (*b*) 15 seconds' incubation with $S^{35}O_4^{--}$ followed by 15 seconds' incubation with $S^{32}$ chaser, (*c*) 15 seconds $S^{35}O_4^{--}$ followed by 120 seconds with chaser. Note transfer of radioactivity from 70–85S region to nonsedimenting region.

1959; Spahr and Hollingworth, 1961) or equally distributed among the 30S and 50S particles (Bolton, 1959).

Deoxyribonuclease (DNase) is also present in ribosomes in a latent form (Elson, 1959).

The significance of other enzymes found in ribosome pellets is uncertain. The pellet is apt to be contaminated by soluble protein and adsorption of enzyme by ribosomes is difficult to rule out.

A major proportion of leucine amino peptidase was found associated with ribosomes of both 30S and 50S. The profile of enzyme activity corresponds to that of ribosomes in sedimentation analysis, ruling out simple contamination of pellets (Fig. 5). This enzyme shows no latency (Bolton and McCarthy, 1959).

Small traces of β-galactosidase continue to sediment with ribosomes after prolonged washing procedures (Fig. 7). Here again the activity of the ribosome associated enzyme is partially latent. Acid phosphatase and alkaline phosphatase have also been found in *E. coli* ribosomes (Cowie *et al.*, 1961).

### 3. *Nascent Protein*

Kinetic studies show the rapid appearance of newly incorporated $S^{35}$- and $C^{14}$-amino acids in ribosomes (Fig. 8). The quantity corresponds to about 0.1% of the total soluble protein, i.e., to the protein synthesized in a period of about 5 seconds (McQuillen *et al.*, 1959). In studies of the synthesis of the structural protein of ribosomes care must be exercised to distinguish the nascent protein, which is only transiently associated with ribosomes and not destined to become a part of their structure.

## D. NUCLEIC ACID COMPONENTS

Degradation of the 50S ribosomes by sodium dodecyl sulfate or by phenol releases RNA of 16 and 23S corresponding to molecular weights of $0.55 \times 10^6$ and $1.1 \times 10^6$. The 30S ribosomes release a single component of 16S (Kurland, 1960). Aronson and McCarthy report similar results and describe further a progressive degradation to small units of 13.1S, 8.8S, and 4.4S which was brought about by heating or removal of magnesium by dialysis. The 4.4S component was also found in the RNA extracted from magnesium starved cells (Aronson and McCarthy, 1961; McCarthy and Aronson, 1961). Further evidence for small subunits of ribosomal RNA shows in X-ray analysis (Timasheff *et al.*, 1961).

The nucleotide composition of ribosomal RNA has been reported by several authors and is given in Table II. Bacteria of different DNA composition show a total ribosomal RNA composition which is quite

TABLE II
Nucleotide Composition of Various RNA's

| Organism | DNA | S-RNA | 70S | 50S | 30S | Reference |
|---|---|---|---|---|---|---|
| *Escherichia coli* | A 24 | A | — | 25.6 | 24.6 | Spahr and Tissières (1959) |
| | T 24 | U | — | 22.1 | 21.0 | |
| | G 26 | G | — | 31.4 | 31.6 | |
| | C 26 | C | — | 20.9 | 22.8 | |
| *Escherichia coli* | A 24 | A 19.7 | — | 26.4 | 24.3 | Bolton (1959) |
| | T 24 | U 17.2 | — | 18.3 | 20.5 | |
| | G 26 | G 34.2 | — | 34.8 | 31.6 | |
| | C 26 | C 29.1 | — | 20.5 | 23.6 | |
| *Escherichia coli* | A 24 | A 20.5 | 25.1 | 25.4 | 24.8 | Midgley (1962) |
| | T 24 | U 16.5 | 20.4 | 19.6 | 21.5 | |
| | G 26 | G 33.2 | 32.6 | 33.5 | 31.0 | |
| | C 26 | C 29.8 | 21.9 | 21.5 | 22.7 | |
| *Bacillus subtilis* | A 29 | A 20.2 | 25.9 | 26.2 | 26.5 | Midgley (1962) |
| | T 29 | U 17.6 | 20.8 | 19.3 | 21.6 | |
| | G 21 | G 32.9 | 31.0 | 32.0 | 29.6 | |
| | C 21 | C 28.3 | 22.3 | 22.5 | 22.3 | |
| *Proteus vulgaris* | A 32 | A 19.1 | 26.2 | 26.5 | 24.7 | Midgley (1962) |
| | T 32 | U 18.3 | 20.7 | 20.8 | 20.4 | |
| | G 18 | G 33.3 | 31.4 | 31.4 | 31.9 | |
| | C 18 | C 29.3 | 21.7 | 21.3 | 23.0 | |
| *Aerobacter aerogenes* | A 22 | A 19.7 | 25.5 | 25.6 | 25.3 | Midgley (1962) |
| | T 22 | U 18.8 | 21.1 | 21.2 | 21.5 | |
| | G 28 | G 32.3 | 31.5 | 31.2 | 30.8 | |
| | C 28 | C 29.2 | 21.9 | 22.0 | 22.4 | |
| *Pseudomonas aeruginosa* | A 18 | A 20.8 | 25.7 | 26.3 | 25.1 | Midgley (1962) |
| | T 18 | U 17.1 | 21.0 | 21.3 | 20.5 | |
| | G 32 | G 33.8 | 31.6 | 31.2 | 32.8 | |
| | C 32 | C 28.3 | 21.7 | 21.2 | 21.6 | |

constant. The increased accuracy obtained by use of the isotope dilution technique shows a slight difference in composition between the RNA's derived from 30S and 50S ribosomes.

## E. STRUCTURE OF RIBOSOMES

Table III lists the values of the physical constant of purified *E. coli* ribosomes as measured by Tissières *et al.* Molecular weights of $0.9 \times 10^6$ and $1.8 \times 10^6$ for the 30S and 50S particles are consistent with the molecular weights of $0.55 \times 10^6$ and $1.1 \times 10^6$ found for the RNA, assuming an RNA content of 63%.

TABLE III
PROPERTIES OF *E. coli* RIBOSOMES[a]

| Particle | S | D | $\bar{V}$ | $\eta$ | MW[b] | MW[c] |
|---|---|---|---|---|---|---|
| 30 | 30.6 | 2.95 | 0.64 | 0.080 | 0.7 | 1.0 |
| 50 | 50.0 | 1.91 | 0.64 | 0.054 | 1.8 | 1.8 |
| 70 | 69.1 | 1.83 | 0.64 | 0.061 | 2.6 | 3.1 |
| 100 | 100.0 | — | 0.64 | 0.071 | — | 5.9 |

Key: S, sedimentation coefficient $S^0_{20,w} \times 10^{13}$ cm/sec; D, diffusion coefficient $D^0_{20,w} \times 10^7$ cm²/sec; $\bar{V}$, partial specific volume; $\eta$, viscosity cm/dl; MW, molecular weight $\times 10^{-6}$.

[a] From Tissières *et al.* (1959).

[b] Calculated from S and D.

[c] Calculated from S and $\eta$.

Electron microscopy of purified ribosome preparations provides the dimensions shown in Table IV (Hall and Slayter, 1959; Huxley and Zubay, 1960).

Electron microscopy of sectioned bacteria is more ambiguous. Some sections show dense regions which may be due to ribosomes, others which might be expected to show ribosomes do not. It is not at all clear whether these differences are due to differences in staining techniques or whether the ribosomes do not exist as compact spheres in the living cell (Hanzon *et al.*, 1959).

TABLE IV
DIMENSIONS OF *E. coli* RIBOSOMES

| Particle | Shape | Dimensions (Å) | Mol. wt. $\times 10^{-6}$ |
|---|---|---|---|
| 30S | Prolate | 95 × 170 | 0.76 |
| 50S | Oblate | 170 × 140 | 2.0 |
| 70S | Oblate | 200 × 170 | 3.4 |

The protein/nucleic acid ratio of 37/63 corresponds closely to a ratio of 2 amino acids/nucleotide. Furthermore, a portion of the protein can be removed (see Section III,A) leaving another portion still in the ribosome structure. Electron microscopy of ribosomes stained with uranyl acetate shows no evidence of a protein shell around a nucleic acid core (Huxley and Zubay, 1960). These findings suggest that the 2/1 ratio may not be fortuitous but a consequence of saturating the entire RNA strand with protein along its entire length.

The hyperchromicity of ribosomes is 40%, the same as that of the RNA after isolation from the ribosomes. Thus, some aspects of the RNA structure are not appreciably altered by the addition of protein (Schlessinger, 1960). A similar conclusion is drawn from X-ray diffraction studies (Zubay and Wilkins, 1960).

## III. Fractionation of Ribosomes

### A. CHROMATOGRAPHY

Chromatography on DEAE cellulose columns provides a means for separating protein, S-RNA, ribosomal RNA, and DNA from ribosomes. It is particularly useful in distinguishing the precursors in ribosome synthesis.

The column (1 $cm^2 \times 10$ cm) is prepared by packing a slurry of DEAE, thoroughly washed with tris buffer (0.01 $M$, pH 7.6, 0.01 $M$ Mg), at 3–5 psi pressure. The preparation to be analyzed (whole cell juice or ribosome pellet) is adsorbed on the column from this buffer and eluted by 200 ml of the buffer in which there is a linearly increasing concentration of NaCl (0–1 $M$).

Under these conditions ribosomes elute at 0.4 $M$ NaCl, S-RNA at 0.5 $M$, and DNA at 0.6 $M$. Ribosomal RNA (prepared by phenol) remains adsorbed. The ribosomes emerge as 30 and 50S particles with their full complement of protein. If the magnesium concentration is

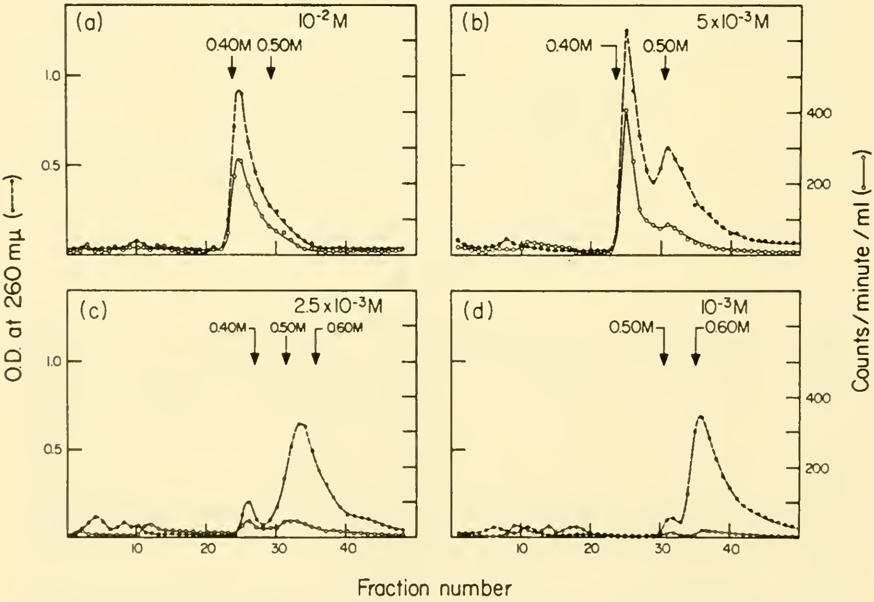


FIG. 9. Chromatography on DEAE of 70S ribosomes using different concentrations of magnesium. Protein content is indicated by $S^{35}$ radioactivity.

decreased the ribosomes lose protein and elute at higher salt concentrations, as shown in Fig. 9.

The use of the DEAE column to separate ribosomes and their precursors for kinetic analysis was first reported in 1957 (Roberts *et al.*, 1958). Recent examples of this technique are given in Section V,A. Separation is achieved because of the different protein content of the ribosome and precursors.

### B. SEDIMENTATION ANALYSIS

The separation of ribosomes of different sedimentation coefficients can be carried out by sedimenting a layer of ribosome suspension through a solution stabilized by a density gradient (Brakke, 1953; Britten and

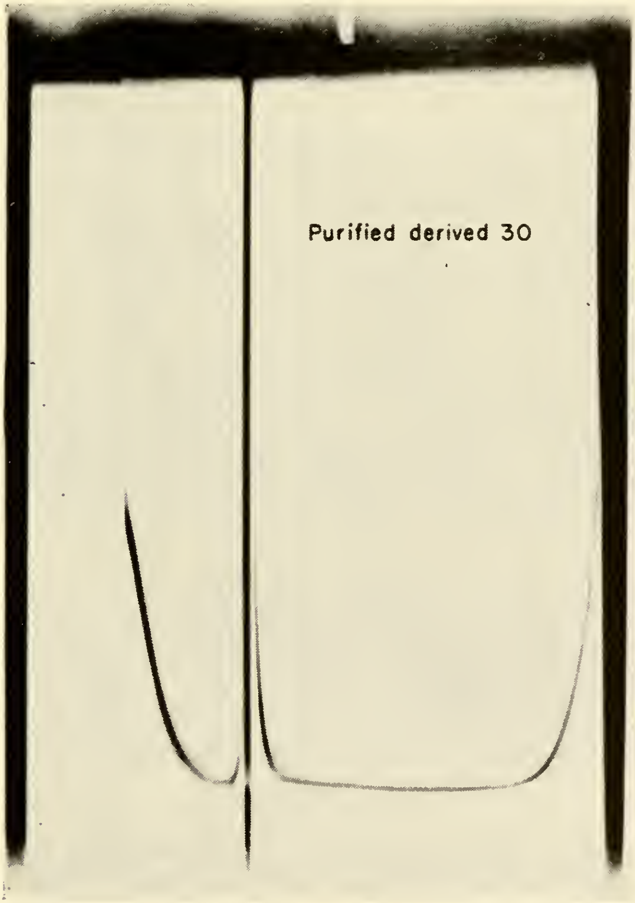


FIG. 10. Analytical centrifuge pattern of 30S ribosomes purified by centrifugation through a sucrose density gradient.

Roberts, 1960). Figure 10 shows the purity of 30S ribosomes prepared in this way. When small quantities of ribosomes are used in very thin layers excellent resolution can be achieved, as shown in Section V,B below.

### C. DENSITY GRADIENT BANDING

Since DNA, RNA, and ribosomes have different densities they can be separated by prolonged centrifugation in appropriate solutions of high density. The materials form bands at the levels where their density equals the density of the solution. Materials differing by only a few per cent in density can be well separated since the density of the solution varies only slightly from top to bottom of the centrifuge tube. Density banding was used to detect the DNA-RNA hybrid complex (Hall and Spiegelman, 1961; Spiegelman *et al.*, 1961) and to show an association between DNA and ribosomes (Nisman, 1961).

Density banding also provides a unique method of separating newly formed ribosomes from old ones. The cells are first grown in media containing heavy isotopes ($C^{13}$, $N^{15}$, and $H^2$) and then transferred to media containing light isotopes. The ribosomes formed before and after the transfer can then be separated by density gradient banding. The value of this technique is illustrated in the work of Brenner *et al.* in their studies of phage-infected cells (Brenner *et al.*, 1961).

### D. ELECTROPHORESIS

Electrophoresis has been used occasionally in the study of ribosomes. Pardee *et al.* showed a difference in the electrophoretic pattern caused by growth of the cells in chloramphenicol (Pardee *et al.*, 1957). Electrophoresis was used to isolate the RNA formed after phage infection (Nomura *et al.*, 1960).

### E. TWO-PHASE SYSTEMS

Successful separation of ribosomes from other cellular components by partition in two-phase systems is described by Albertsson (1960). This method has not been exploited extensively in the study of ribosomes but it gives promise of being extremely useful in the search for complexes of ribosomes with DNA or for other studies of cellular organization.

### F. CONCLUSION

There is still no general technique for distinguishing ribosomes of slightly different composition. Hence there is no method for determining whether there are two or a thousand different types of ribosomes. The only technique applied to date is the specific precipitation of the ribosomes which carry a particular protein by the antibody to that protein. Ribosomes have been precipitated by specific antibodies (Cowie *et al.*, 1961; Warren and Goldthwait, 1961) but analysis shows only a marginal difference in base composition from the average (Midgley and Goldthwait, 1961).

## IV. METABOLIC POOLS

For kinetic studies of ribosome synthesis to be interpretable in a quantitative way a number of conditions must be met. Among these, two are of primary importance. First, the cells must be growing in steady-state conditions so that cellular components are not changing in their proportion. Second, the time course of the specific radioactivity of the immediate precursors for macromolecular synthesis must be known.

The first condition can be satisfied by allowing a prolonged period of exponential growth (in an adequate and unchanging medium) before the introduction of the tracer. The addition of a tracer such as $S^{35}O_4^{--}$ or $P^{32}O_4^{---}$ to a medium which already contains $S^{32}O_4^{--}$ and $P^{31}O_4^{---}$ introduces no disturbance. In principle, labeled amino acids or bases should be added to a medium which already contains the unlabeled amino acid or base without changing the concentration. In practice, however, this procedure is wasteful of the tracer, and no transient in the rate of macromolecular synthesis occurs as a result of the change from an endogenous to an exogenous supply of amino acids or bases. The customary procedure is therefore to add a labeled amino acid or base to a culture already growing exponentially on its endogenous supply. Deficient mutants should not be used in this way as the addition of the *required* (labeled) nutrient to a starved cell would initiate growth, steady-state conditions would not apply, and the results would not be meaningful.

The second condition requires a technique for the evaluation of the specific radioactivity of the precursors to the macromolecules. In *E. coli* there exist pools of amino acids in the cell so that exogenous tracer amino acids may be diluted by endogenous amino acids upon entry. Also there is exchange between exogenous and endogenous amino acids so that even the external material can be diluted. Following the addition of an amino acid to the medium, it is concentrated (as much as 5000 times) in the cells and depleted from the medium. Furthermore, an amino acid may be converted to another before incorporation into protein. These reactions would be difficult to follow in detail by any direct method.

Fortunately, when steady-state conditions prevail, the required information can be derived indirectly from simple measurements. The rate of protein synthesis per unit mass of cells is constant; thus, the total rate of entry of radioactivity into proteins of all types must measure the specific radioactivity of the precursor. Direct measurements of the pool and product protein have shown that the rate of entry into the protein was proportional to the specific radioactivity of the pool. This observation indicated that the amino acid pool was in fact a precursor to protein (Britten *et al.*, 1955).

The rapid establishment of the full rate of entry of $S^{35}O_4^{--}$ into protein indicated that there was no appreciable pool of sulfur compounds (McQuillen *et al.*, 1959). Leucine also shows prompt entry into protein (Britten *et al.*, 1962) but a mixture of amino acids from a protein hydrolyzate may show a lag before the full rate is established. This is presumably due to the existence of quite large pools of glutamic acid, alanine, and aspartic acid.

$P^{32}O_4^{---}$ is far from an ideal tracer because of the large pool of phosphate and phosphorylated compounds. There is a further delay introduced by the time required for the $P^{32}$ to reach the $\alpha$ position of the nucleotides (Bolton and Roberts, 1956). In *E. coli* the specific radioactivity of the nucleic acid precursors rises roughly as $(1 - e^{-t/T_0})$ where $T_0$ is 600 seconds. This slowly rising term makes it very difficult to observe the kinetic delays introduced by subsequent smaller pools of ribosome precursors.

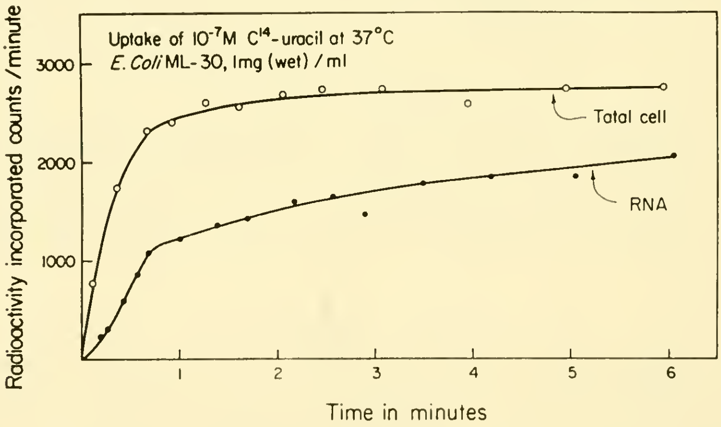


FIG. 11. Incorporation of $C^{14}$-uracil into the metabolic pool and the RNA of *E. coli*. The difference between the two curves measures the radioactivity of the pool. The initial concentration was low; $10^{-7}$ $M$.

The nucleic acid bases provide much more suitable tracers for kinetic studies of RNA synthesis. In spite of the presence of pools of nucleoside di- and triphosphates, exogenous purines and pyrimidines are rapidly incorporated into nucleic acid (Bolton and Roberts, 1956). Recent studies of incorporation of $C^{14}$-uracil (McCarthy and Britten, 1962) and other bases (Buchwald and Britten, 1962) have given a much more detailed understanding of the process.

Figures 11 and 12 show two characteristic features of the incorporation process. At low concentration two phases of the process are evident, one a rapid rate of incorporation into nucleic acid which persists as long as external uracil is present, followed by a much reduced rate which occurs when the uracil is utilized solely from the pool. At higher concentrations the two phases are also observable. However, in this case the rate of incorporation during the second phase is more rapid, as the direct incorporation continues throughout and is supplemented during the second phase by $C^{14}$-uracil drawn from the pool.

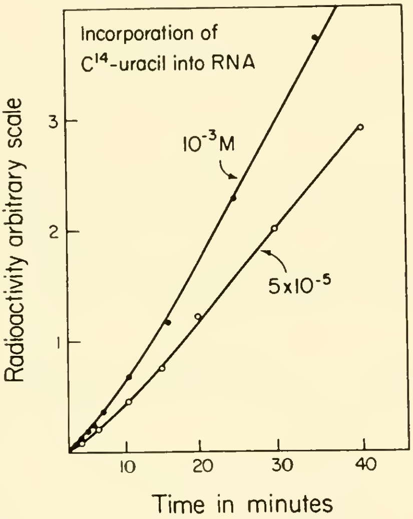
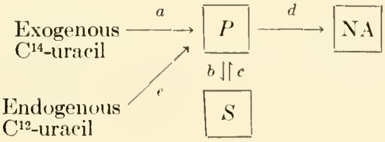


FIG. 12. Incorporation of $C^{14}$-uracil into the RNA of *E. coli* from high concentrations of external uracil. The ordinate scales of the two curves are different.

To account for the direct entry of $C^{14}$-uracil together with utilization of pool material a mechanism such as that shown in the diagram below must be envisaged.



The maximum entry flow ($a$) is equal to the flow ($d$) which is required for the synthesis of nucleic acid (NA). Furthermore, the flow due to endogenous synthesis ($e$) is regulated by feedback control so that $a + e = d$. Thus the pool ($S$) remains constant and cannot be expanded or depleted. The pool ($P$) cannot contain more than 5 seconds' supply of uracil compounds or an observable kinetic delay would appear in the entry of $C^{14}$-uracil into NA. Forty per cent of the exogenous uracil proceeds directly to NA and 60% enters the pool $S$, so that the flow ($b$) is roughly 60% of ($a$).

Labeled bases, once incorporated into the pool, are not rapidly removed by exchange with external material. Hence the incorporation of tracers into nucleic acid cannot be abruptly terminated by dilution of the exogenous tracer material. The best condition which can be achieved is the abrupt transition from phase 1 to phase 2. By allowing only a very short time before the external material is exhausted, a ratio of 20:1 in the rates of phase 1 and phase 2 can be obtained.

The conversion of uracil to cytosine causes a further complication in the kinetics of incorporation because the proportion of the labeled material which enters nucleic acid as cytosine increases with time.

Guanine, adenine, and cytosine show a similar behavior but the pool sizes and the proportions entering directly differ, as shown in Table V.

TABLE V

INCORPORATION CHARACTERISTICS OF FOUR RNA BASES

| Base | Direct entry (%)[a] | Pool ($S$) time constant (minutes)[b] |
|---|---|---|
| Uracil | 40 ± 5 | 10 ± 2 |
| Cytosine | 45 ± 5 | 28 ± 5 |
| Guanine | 78 ± 5 | 3.3 ± 0.5 |
| Adenine | >50 | 3–9[c] |

[a] Shows the initial rate of incorporation of the tracer into RNA as per cent of the ultimate rate, per unit cell mass.

[b] Shows the time constant ($1/e$) of the pool ($S$) derived from the exponential decay of its radioactivity during the second phase of an experiment at low concentration such as Fig. 11. These values are in agreement within the errors indicated with the values derived from the time constant with which the incorporation into RNA (at high concentrations such as Fig. 12) approaches its ultimate rate.

[c] The pool of adenine compounds does not show a single time constant and a detailed analysis to resolve the individual time constants has not been carried out.

Once the complications introduced by the pool are recognized and measured, appropriate corrections to the kinetics observed in the macromolecular components can be made. Without this understanding, the two phases of incorporation into nucleic acids might be misinterpreted as indicating an exchanging component of the nucleic acid.

## V. Kinetics of Incorporation into RNA of Ribosomes

Kinetic studies of ribosome synthesis depend upon methods of separating various classes of macromolecules. The techniques which have proved to be most useful are chromatography and sedimentation analysis. The latter technique can be used both for the analysis of the ribonucleoprotein particles and of the RNA extracted from them.

In the following paragraphs details have been drawn exclusively from the work carried out in our laboratory (Roberts *et al.*, 1958; McCarthy and Aronson, 1961; McCarthy and Britten, 1962; Britten *et al.*, 1962). The reasons for this are twofold. On the one hand, we are more familiar with our own work, and on the other, different experiments can be related more surely because they were carried out under similar conditions. Experiments in other laboratories also contribute to our understanding.

An RNA fraction of high specific radioactivity has been separated by electrophoresis (Nomura *et al.*, 1960). Otaka *et al.* have observed precursor product relationships among fractions separated on DEAE (Otaka *et al.*, 1962). Gros *et al.* have studied an RNA component of high specific radioactivity which is dissociated from ribosomes when the magnesium concentration is reduced and which has an average sedimentation coefficient of about 12S (Gros *et al.*, 1961).

### A. DEAE chromatography

Figure 13 shows the elution patterns obtained from cells given progressively longer exposures to $C^{14}$-uracil. After breaking the cells, the walls and membranes were removed by a brief period of centrifugation. Ribosomes and their precursors were sedimented by prolonged centrifugation (40K 240′) leaving the S-RNA in the supernatant fluid. This procedure is desirable because S-RNA elutes at 0.5 *M* NaCl contaminating ribosome precursors. DNA, which also elutes in this region, was removed by DNase.

The patterns show clearly that the $C^{14}$-uracil first appears in material eluting at 0.6 *M* NaCl. After the first few minutes this component reaches a saturation level and $C^{14}$-uracil builds up in material eluting at 0.5 *M* NaCl. As time goes on the second region approaches saturation and $C^{14}$-uracil appears in the bulk of the ribosomes which elute at 0.4 *M* NaCl. The column seems to have separated two sequential precursors from the final product. This qualitative impression is confirmed by the

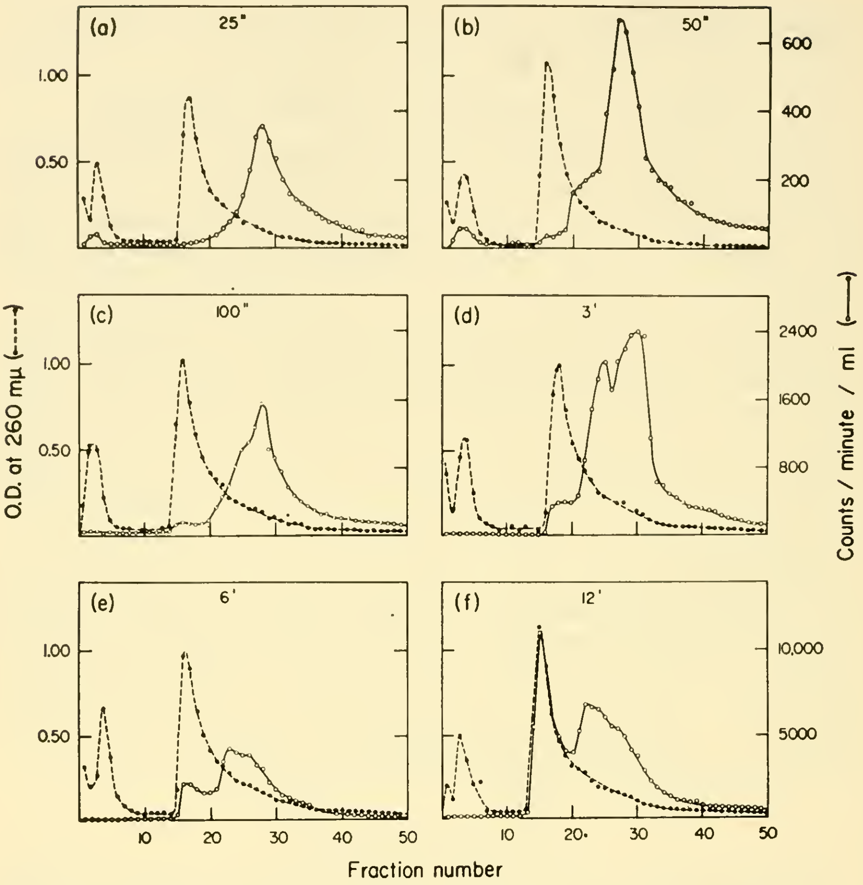


Fig. 13. Analysis of total ribosomal pellets (40K 240 minute pellets) on a DEAE-cellulose column. Linear sodium chloride gradient from 0.2 *M* to 1.2 *M* in tris-HCl buffer 0.01 *M* containing $MgCl_2$ 0.01 *M*, pH 7.4. Salt gradient 0.004 *M*/ml. Volumes collected 3.6–3.8 ml. (*a*) 25 seconds exposure to $C^{14}$-uracil, (*b*) 50 seconds, (*c*) 100 seconds, (*d*) 3 minutes, (*e*) 6 minutes, (*f*) 12 minutes. The three sequentially labeled peaks appear at 0.6 *M*, 0.5 *M*, and 0.4 *M* NaCl.

quantitative analysis given below (Section V,D). As these two precursors have quite distinct elution properties, as well as other characteristics to be described later, we have used the names "eosome" and "neosome" to distinguish them from ribosomes and to indicate their order in the sequence of synthesis.

### B. SEDIMENTATION ANALYSIS

Most of the earlier studies of the kinetics of ribosome synthesis were carried out by means of sedimentation analysis in a magnesium concentration of $10^{-2}\,M$, which is necessary for the preservation of large ribosomes. At early times the specific radioactivities of the 30S and 50S ribosomes present in the extract were considerably higher than that of the 70S. These small particles, described as native 30S and 50S to differentiate them from those derived by breakdown of the 70S at lower magnesium concentrations, represent together some 10 to 20% of the ribosomal material. Their higher specific activity after a pulse was interpreted as showing a precursor relationship to the 70S ribosomes.

Figure 14 shows the analysis of extracts of four samples of cells taken after 10, 20, 50, and 55 minutes' exposure to $C^{14}$-uracil. 50S and 30S ribosomes show as shoulders in the ultraviolet absorption profile. At 10 minutes the specific radioactivities of the 50S and 30S ribosomes are at least three times that of the 70S. This difference in specific radioactivity decreases with time and has practically disappeared by 55 minutes.

When a series of cell extracts was analyzed after much shorter exposures to $C^{14}$-uracil the results were not interpretable in terms of a simple hypothesis. At early times an appreciable fraction of the radioactivity appeared in the 70S particles. The quantity was greater than would be expected in an end product. This suggested direct entry of RNA into 70S without the delay brought about by passing through the large pools of 50S and 30S precursors. Another possibility was that the labeled RNA was loosely associated with the 70S peak rather than being incorporated into the large ribosomes.

Because of the complexity encountered in the analysis of cell extracts made in $10^{-2}\,M$ $Mg^{++}$, a search was made for conditions more suitable for the resolution of precursors from products in ribosome synthesis. Comparison was made of a pulse labeled cell extract prepared and analyzed in a range of magnesium concentrations (Fig. 15).

At $10^{-2}\,M$ $Mg^{++}$ (Fig. 15a) most of the ribosomes are present as 70S but most of the radioactivity runs as 50S and 30S. The apparent specific radioactivity of the small ribosomes is more than three times that of the 70S. Cells broken in the presence of $3 \times 10^{-3}\,M$ $Mg^{++}$ (Fig. 15b) have most of their ribosomes as 50S and 30S. At the same time about half of the radioactivity associated with ribosomes in Fig. 15a now appears in a peak of about 14S. There is now little difference between the specific radioactivities of the three groups of ribosomes although there is a peak of radioactivity running between the main 50S and 30S peaks. In $10^{-3}\,M$ magnesium (Fig. 15c) an even higher proportion of the radioactivity appears in a 14S peak, and the faster of the other two peaks of radioactivity appears to lag behind the 50S ribosome peak. The distribution of radioactivity between the three objects is very similar at $10^{-4}\,M$ $Mg^{++}$

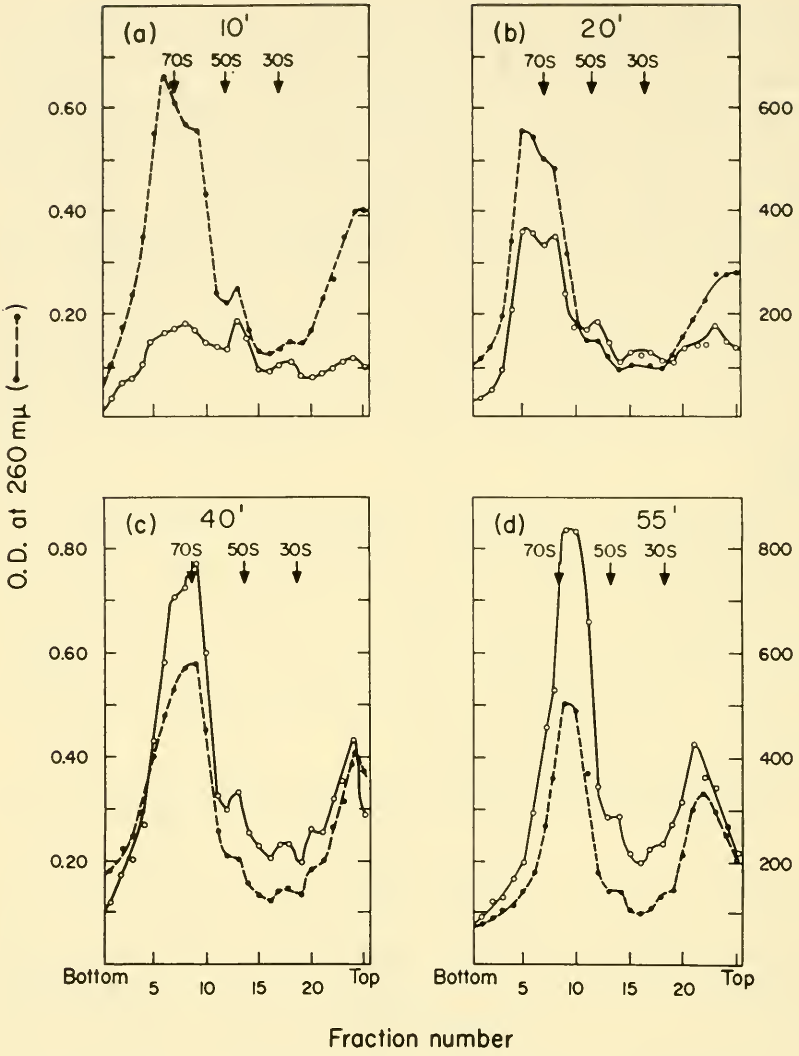


FIG. 14. Sedimentation analysis of four total cell extracts taken from cells given (*a*) 10-minute, (*b*) 20-minute, (*c*) 40-minute, and (*d*) 55-minute exposures to $C^{14}$-uracil. Centrifugation 90 minutes at 37,000 rpm. Cell washing and breakage and centrifugation carried out in tris-HCl 0.01 *M*, pH 7.4, $MgCl_2$ 0.01 *M*.

but the increased centrifugation time in Fig. 15d produces higher resolution. In this instance the front-running peak of radioactivity is clearly resolved from 50S ribosomes and appears to have a sedimentation coefficient of about 43S.

The resolution of early labeled RNA from the bulk ribosomes sug-

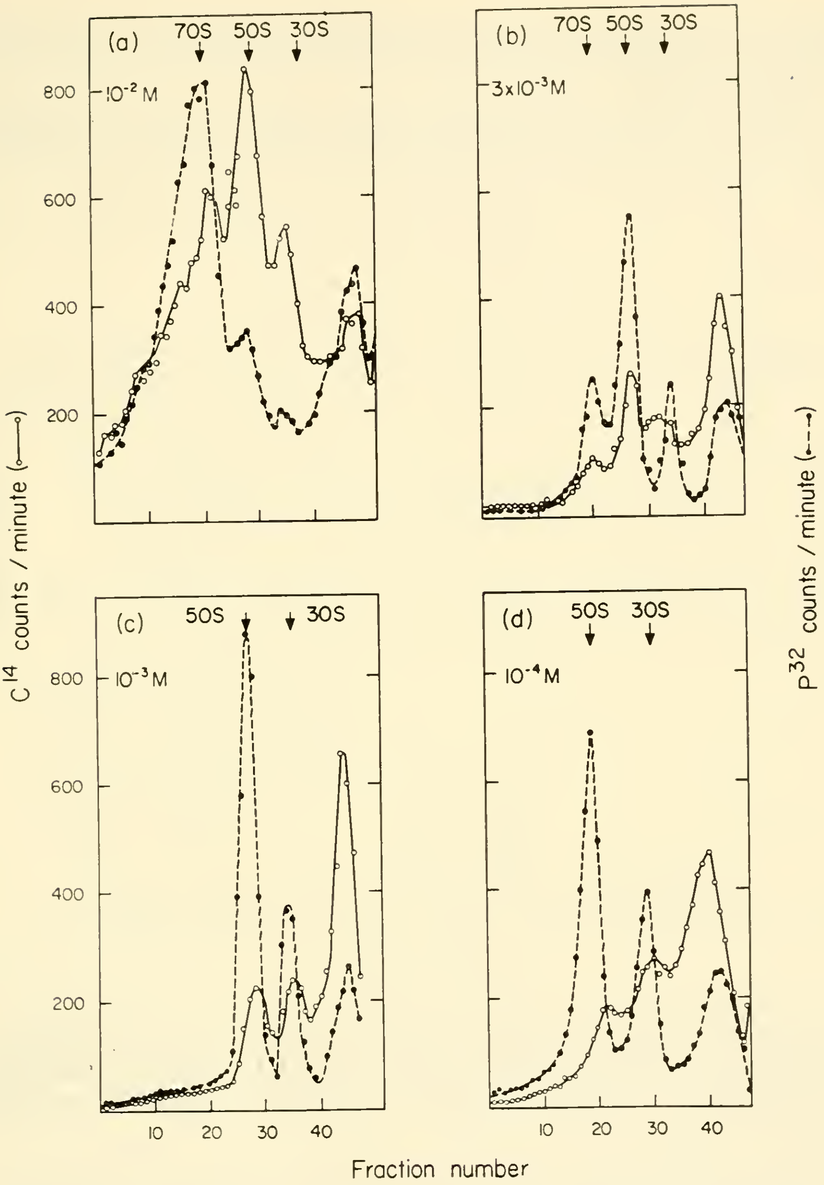


FIG. 15. Sedimentation analysis of a total extract from cells labeled with $P^{32}$ for four generations and with $C^{14}$-uracil for 2 minutes. Cells washed and extracts prepared and run in tris-HCl 0.01 *M*, pH 7.4 (*a*) containing $MgCl_2$ at $10^{-2}$ *M*, (*b*) $3 \times 10^{-3}$ *M*, (*c*) $10^{-3}$ *M*, (*d*) $10^{-4}$ *M*. Centrifugation: (*a*), (*b*), and (*c*) 90 minutes at 37,000 rpm; (*d*) 150 minutes at 37,000 rpm, 4°C.

gested that these conditions of cell breakage and analysis of extracts would be suitable for the determination of the details of the kinetics of ribosome synthesis.

Therefore, analyses were made in $10^{-4}$ $M$ magnesium of six samples of extracts prepared from $P^{32}$ steady-state labeled cells given from 30 seconds to 12 minutes' exposure to $C^{14}$-uracil (Fig. 16). The extracts were treated with DNase prior to centrifugation. The $P^{32}$ profile shows three main peaks of 50S and 30S ribosomes and soluble RNA.

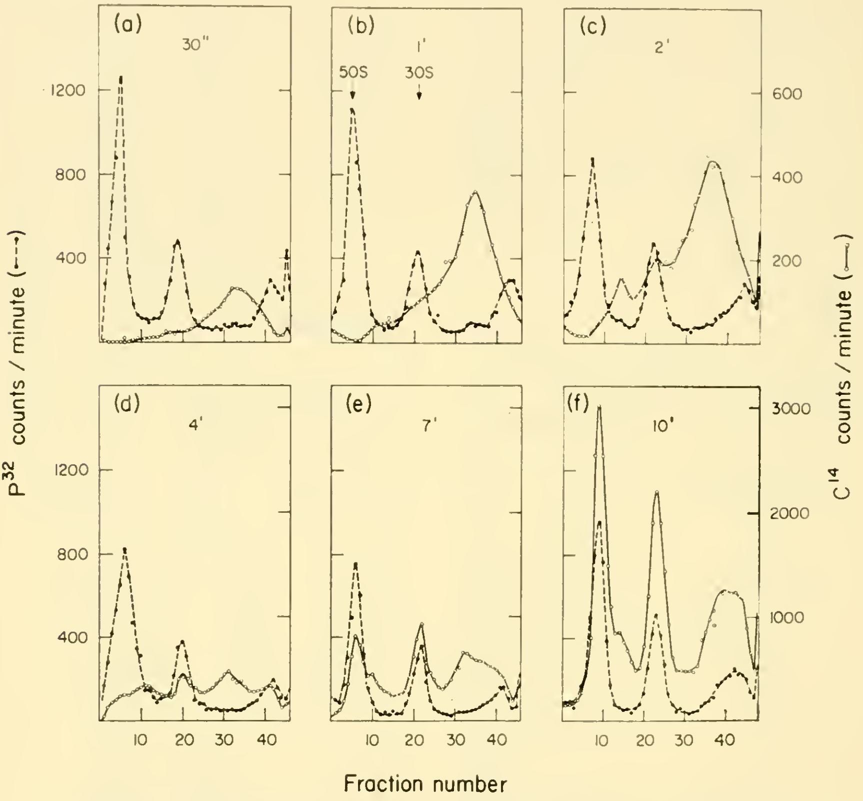


FIG. 16. Sedimentation analysis of six DNase treated total extracts made and centrifuged in tris-HCl 0.01 $M$, pH 7.4, $MgCl_2$ $10^{-4}$ $M$ from cells steady state labeled with $P^{32}$ given (*a*) 30 seconds' exposure to $C^{14}$-uracil, (*b*) 1 minute, (*c*) 2 minutes, (*d*) 4 minutes, (*e*) 7 minutes, (*f*) 10 minutes. Centrifugation: 175 minutes at 37,000 rpm, 4°C.

The patterns indicate a sequence of precursors and products. At the earliest time the $C^{14}$-uracil is carried by material sedimenting in a broad band centered at about 14S. Later, $C^{14}$-uracil shows at 30S and at 43S.

Finally, the $C^{14}$-uracil accumulates in the end products, the 30S and 50S ribosomes. In addition, radioactivity builds up in S-RNA which sediments at 4S. A qualitative comparison with the sequence of events revealed by DEAE chromatography suggests that the first precursor, eosome, sediments at 14S and that neosomes are present in both the 43S and the 30S regions. The 30S and 50S ribosomes correspond to the main ribonucleoprotein peak of the column analysis.

### C. QUANTITATIVE KINETIC ANALYSIS

To compare the results of experiments with theoretical models the experimental data are expressed in terms of functions $\phi_X$. These functions and the other needed symbols are defined below.

$\tau$: Time after addition of tracer; units such that $\tau = 1$ when the cells have grown by a factor $e$; $Q = Q_0e^\tau$ gives the growth of the cells or any components.
$\mu$: Effective specific radioactivity of tracer; units such that the specific radioactivity of the RNA will approach $\mu$ after a long period of growth at a constant tracer concentration.
$X$: Quantity of a component, in general.
$X^*$: Its radioactivity.
$\mu_X$: Its specific radioactivity.
$E, N, R$: Quantities of three specific components.
$T$: The total of all components.

$$\phi_X = \frac{\mu_X}{\mu} \cdot \frac{X}{T} = \frac{X^*}{\mu T} \tag{1}$$

This choice of units is designed to allow different experiments carried out with different growth rates and different effective tracer radioactivities to be plotted on the same graph.

While $\phi_X$ is adequately defined by Eq. (1), the following general definition may perhaps be useful. $\phi_X$ is the ratio of the number of labeled atoms at a given time in a given class of molecules to the number of atoms in the whole cell which would be labeled after a long period of growth during which the labeling conditions were precisely constant. $\phi_X$ is thus a measure of the number of newly synthesized molecules which occur in a given class, and has been named the "newly" synthesized fraction."

When there are two sequential precursors ($E$ and $N$) to a product $R$

$$\rightarrow \boxed{E} \rightarrow \boxed{N} \rightarrow \boxed{R}$$

$$\phi_T = 1 - e^{-\tau} \tag{2}$$

$$\phi_E = \frac{E}{T}\{1 - \exp[-(T/E)\tau]\} \tag{3}$$

$$\phi_N = \frac{N}{T}\left\{1 - \frac{E(R+N)\exp[-(T/E)\tau] - TN\exp[-(R/N)\tau]}{E(R+N) - TN}\right\} \tag{4}$$

and

$$\phi_R = \phi_T - \phi_E - \phi_N \tag{5}$$

The derivation of these equations and the restrictive conditions governing their use are given in (Britten and McCarthy, 1962).

It can be seen from Eq. (1) that the ultimate value of $\phi$ for any

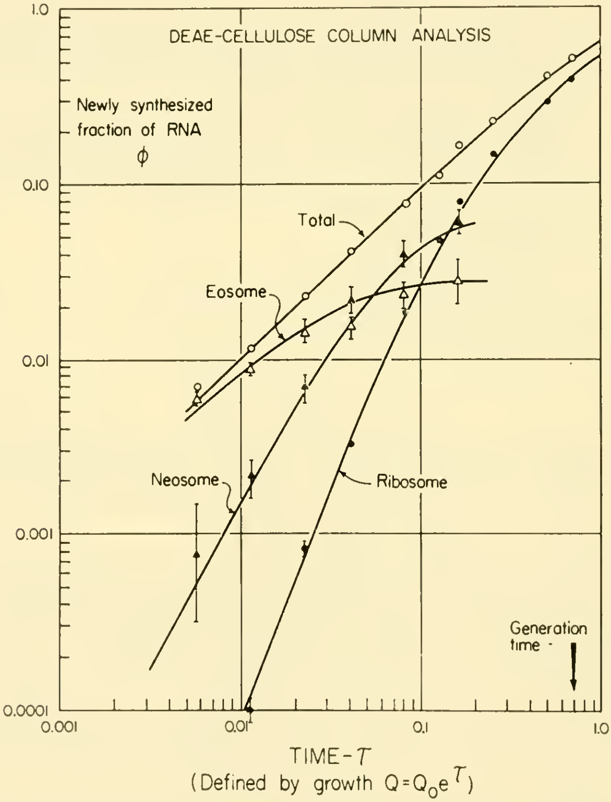


FIG. 17. Log-log plot of the time course of the radioactivity (expressed as $\phi$) of the ribonucleoprotein fractions resolved by DEAE chromatography. Data from Fig. 13 and Fig. 1 of McCarthy *et al.* (1962). The four curves have been calculated from Eqs. (2), (3), (4), and (5) for the case $E = 0.027$, $N = 0.068$, $R = 0.905$.

component is the proportion of that component. At early times both $\phi_E$ and $\phi_T$ increase as $\tau$, $\phi_N$ increases at $\tau^2$ and $\phi_R$ increases as $\tau^3$.

In Fig. 17, $\phi_N$ is plotted for the two precursor fractions and the product resolved by chromatography, together with theoretical curves calculated from Eqs. (2), (3), and (4) for $E = 2.7\%$, $N = 6.8\%$. (The details of the procedure for deriving values of $\phi$ from the experimental data are described in (McCarthy *et al.*, 1962). The quantitative analysis thus verifies the assumption that the three components show the kinetic features of two sequential precursors and a stable product.

The corresponding analysis of the peak resolved by sedimentation is more complicated since there are more components to consider. As a first step the 14S region is compared with other regions of the sedimentation pattern by plotting its radioactivity ($\phi_E$) together with $\phi_T$ and $\phi_{R'}$ ($\phi_{R'} = \phi_{50} + \phi_{43} + \phi_{30}$) (Fig. 18). Initially, all of the radioactivity

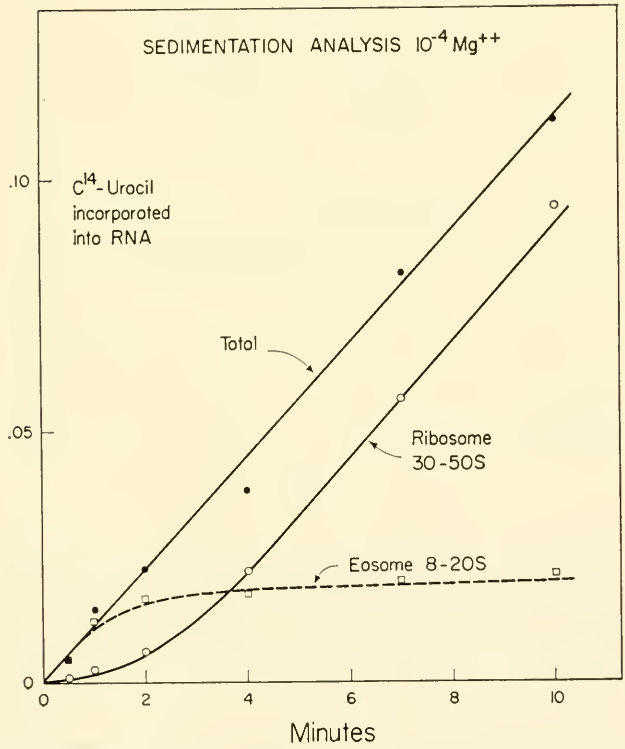


FIG. 18. The $C^{14}$-uracil label incorporated into total RNA (●), the eosome region 8–20S (□), and the ribosome region (○), i.e., material between 30S and 50S including neosome and ribosome. Data from Fig. 16. The label in each component is plotted as $\phi$ functions $\phi_T$, $\phi_E$, $\phi_{N+R}$:

appears in the 14S region and the shape of the curve indicates that the total flow of material to ribosomal RNA passes through this stage. The curve for $\phi_{R'}$, has zero slope at zero time showing that the entry of radioactivity is delayed by a precursor and that little, if any, $C^{14}$-uracil by-passes the precursor. The values of $\phi_{R'}$ at early times indicate a precursor quantity of 3.5%. The quantity of precursor estimated from its final level is 2%. Although this discrepancy may indicate complexity of the first precursor, the material which sediments in the 14S region can be identified with the eosome fraction resolved by chromatography.

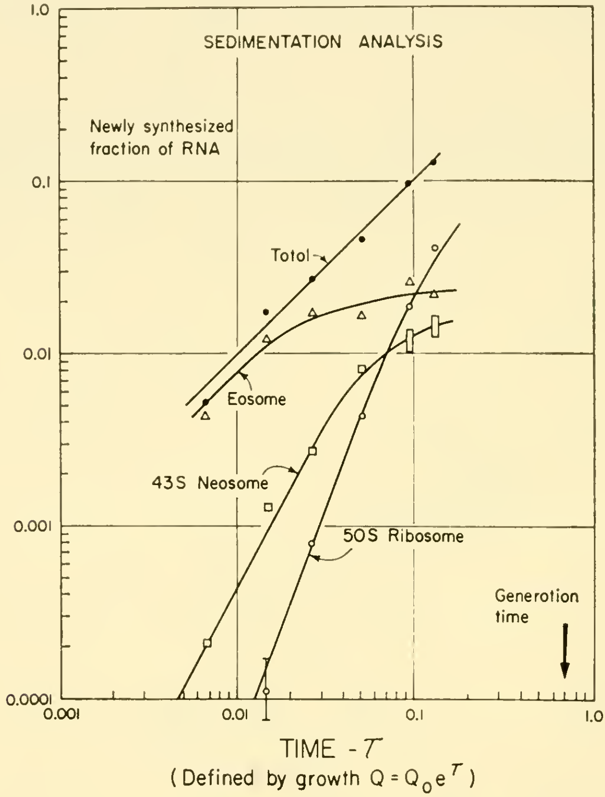


FIG. 19. Log-log plot of the newly synthesized fraction of RNA present in the total cell extract (●), eosome (△), 43S neosome (□), and 50S ribosome (○) as a function of time. Data from Fig. 16. The curve drawn for the 50S ribosomes is the calculated curve shown on Fig. 17 multiplied by 2/3 since the 50S ribosomes account for 2/3 of the ribosomal nucleoprotein. The agreement is striking.

Figure 19 shows $\phi_E$ plotted together with $\phi_{43}$, $\phi_{50}$ and $\phi_T$. At early times $\phi_{43}$ is proportional to $\tau^2$ and $\phi_{50}$ is proportional to $\tau^3$. At late times

$\phi_{43}$ accounts for 2/3 of the neosome fraction resolved by chromatography and $\phi_{50}$ accounts for 2/3 of the ribosome fraction.

The remaining 1/3 of the ribosome fraction is clearly accounted for by the 30S peak. However, this component must also include as an unresolved peak the remaining 1/3 of the neosome fraction. In Fig. 20 $\phi_{30}$ is compared to $\phi_{43+50}$. If the 30S and 50S ribosomes were synthesized independently from a common pool of eosomes, and if the 30S peak included both the neosome and the ribosome stage of synthesis, then $\phi_{43+50}$ should be twice $\phi_{30}$ at all times. At early times, however, the 30S component has more than its share of radioactivity. Furthermore, at

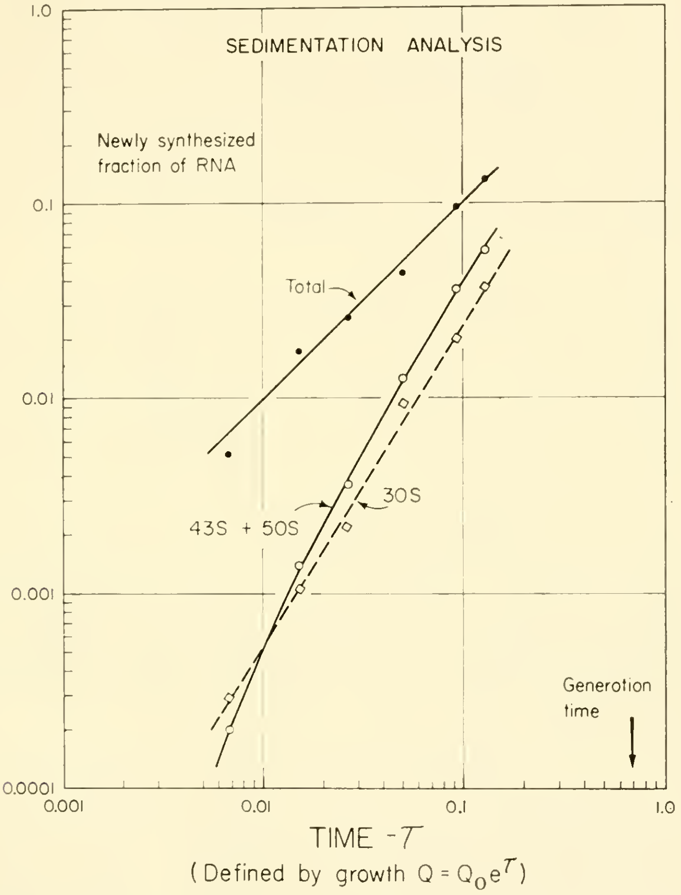


FIG. 20. Log-log plot of the newly synthesized fraction of RNA present in total cell extract (●), 30S region (□), and 43S + 50S region (○) as a function of time. Data from Fig. 16.

early times the 43S component accounted only for 1/3 of the neosome fraction. One interpretation of these results is that the 30S peak includes not only the 30S neosome stage of the 30S ribosome but also a 30S neosome precursor to the 43S.

### D. STAGES OF SYNTHESIS OF THE RIBOSOMAL RNA

From the analysis of uracil-labeled cell extracts, it is clear that at early times all the radioactivity is present in a fraction, named eosome, peaking at 14S. The kinetic behavior of the component has been studied and its size estimated at 2–3% of the total ribosomal RNA. It can be shown from the same experiments that label entering nucleoprotein of higher molecular weight is delayed by a component of about the same size. Together these observations suggest that the eosome is predominantly precursor to ribosomes.

In another recent study (Gros *et al.*, 1961) the rapidity of labeling of the eosome has been invoked as a criterion of turnover. These authors have suggested that the rapid rate with which tracer enters and leaves the eosome fraction proves instability of these RNA molecules. In general, however, it is clear that the rapid labeling of a small component indicates merely that the flow into it is much larger than that required to maintain its quantity in the growing cell. The loss of label when an excess of unlabeled isotope is added shows only the already obvious fact that there is also a transfer of material out of this component. These measurements cannot be expected to distinguish between a precursor through which label flows to a product and a compound which is synthesized and then is broken down to its constituent parts. What we have demonstrated is that there exists a precursor-product relationship between the eosome and the product ribosomes for the transfer of $C^{14}$-uracil radioactivity.

This does not prove that labeled molecules initially observed as eosomes are incorporated *as such* into completed ribosomes. An alternative involves breakdown of the eosome molecules and a synthesis of ribosomes by means of a quantitative reutilization of the labeled degradation products. It should be pointed out that, in common with many similar tracer studies, the latter alternative cannot be distinguished. For quantitative reutilization the breakdown has to occur in such a way that there is no mixing with other RNA precursors. The kinetics of tracer flow are most simply interpreted on the basis of precursor product relationships, although other evidence discussed in Section VIII points to greater complexity. Here we follow the simple interpretation as a first approximation.

As far as the details of ribosome synthesis are concerned, it is clear that three main stages may be described. The clearest demonstration derives from the rate of labeling of the main ribonucleoprotein peak in the column analysis. Its radioactivity rises initially as the cube of time. The total label present at early times is so low as to require two sequential precursors. The total precursor material can be only 10% of the total RNA and a single precursor of this magnitude could not introduce such a delay into the product. The diagram shown in Fig. 21 summarizes the results obtained from chromatography and sedimentation and indicates the correlation between the two analyses. In the following paragraphs the evidence which supports each of the features of this diagram will be considered.

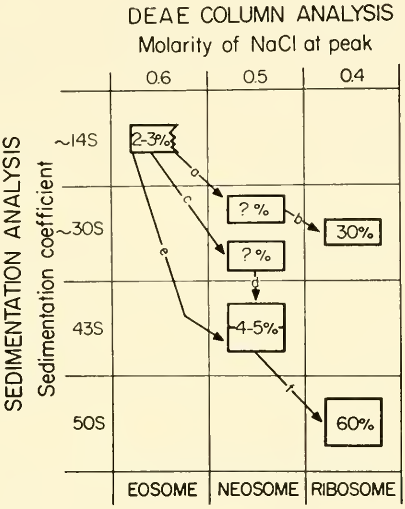


FIG. 21. Schematic diagram of the flow of RNA in ribosomal synthesis. The diagram indicates the way in which the information derived from DEAE-cellulose column analysis and sedimentation analysis has been combined to give a more complete scheme. The figures in the boxes represent the percentages of the total ribosomal RNA existing in each of the states in a steadily growing cell. Movement downward is a result, principally, of the completion of RNA subunits. Movement to the right is principally a result of the addition of protein.

### 1. *Eosome*

The interpretation of the role of the eosome as precursor to all of the ribosomal RNA has been discussed above. Both the peak at 0.6 $M$ in the column analysis and the broad region peaking at 14S in the sedimentation analysis carry all of the radioactivity at early times and later saturate at about the same quantity of radioactivity.

### 2. *Flows a and b and the 30S Neosome Precursor to 30S Ribosomes*

It is clear from the column analysis that there exists a neosome precursor to 30S ribosomes since the peak eluted at 0.4 *M* contains both the 30S and 50S ribosomes, the radioactivity entering it is very strongly delayed and rises as $\tau^3$ at early times. The only location in the sedimentation analysis (Fig. 16) where such an object can occur is in the 30S region. Because of the lack of resolution it has been impossible independently to assess its kinetics of labeling or measure its quantity. After correction for the trails of 43S and eosome peaks in Fig. 16c and d, it appears that the $C^{14}$ radioactivity in the 30S region reaches its maximum somewhat behind the $P^{32}$ peak corresponding to the 30S ribosomes. This result is quite uncertain but leaves an impression of heterogeneity in the 30S region.

### 3. *Flows c and d and the 30S Precursor to 50S Ribosomes*

Figure 20 shows that there is a greater flow into the 30S region than is required for the synthesis of the 30S ribosomes; in fact, more than half the flow from the eosome passes into the 30S region at early times. This shows that a part of the flow that ultimately reaches the 50S ribosome passes through a 30S neosome. We have chosen to indicate on the diagram that ½ of the flow to 50S passes this way for the following reasons. In the first place, the 43S radioactivity (Fig. 19) is 1/3 of the neosome radioactivity (Fig. 17) at early times. In the second place, a sedimentation coefficient of 30S suggests that ½ of the 50S RNA is already present. It is not clear whether the two 30S neosomes indicated are identical to each other. The nucleotide compositions of the 50S and 30S ribosomes differ but this difference could be made up through flow *e*. The total quantity of the 30S neosomes can be crudely estimated by the difference between the amount of 43S neosome (4–5%) and the total amount of neosome (7%) indicated on Fig. 17. No estimate of the relative quantities of the two objects indicated can be made if in fact they differ.

### 4. *The 43S Neosome and Flows e and f*

The 43S neosome stands out clearly in sedimentation analysis (Fig. 16) at the appropriate time. The time course of labeling shown on Fig. 19 is that of the neosome or second stage in the sequence. The radioactivity rises at $\tau^2$ initially and levels off later. At early times the curve has the shape of $\phi_N$ (Fig. 17) but only ⅓ of the magnitude of $\phi_N$. If the total flow to the 50S ribosomes passed directly to the 43S neosome from the eosome, $\phi_{43}$ would be expected to be just ⅔ of $\phi_N$. Therefore, it is clear that about half the flow to 50S ribosomes passes to the 43S neosome from the eosome.

It appears certain that no eosomal RNA goes directly to the 50S ribosome because $\phi_{50}$ is proportional to $\tau^3$ at early times. The 43S neosome is shown on the diagram (Fig. 21) as containing the full complement of RNA of the 50S ribosome. This seems likely from its sedimentation constant.

The quantity of 43S neosomal RNA can be estimated to be about 5% from the steady-state $P^{32}$ radioactivity that remains in this region after the 30S and 50S contributions have been subtracted, assuming reasonable and symmetrical peak shapes. The specific radioactivity estimated on this basis at early times is just ½ what would be expected if the total flow to 50S ribosomes passed directly from the neosome to the 43S. This is, of course, consistent with the diagram since ½ of the flow should be delayed by the 30S neosome precursor to the 43S.

The diagram (Fig. 21) shows a set of sequential relationships which are more complex than the two sequential precursors shown on Fig. 17. The diagram suggests that half of the radioactivity of the 50S ribosomes should rise in proportion to $\tau^4$ for a time while the other half should rise as $\tau^3$. All of the 30S ribosome radioactivity should rise as $\tau^3$.

While deviations from the three-stage model showed clearly in the time course of labeling precursors such as the 30S and 43S neosomes, the accuracy of the data is not adequate to resolve the predicted fourth power component in the final product.

These studies of the synthesis of the RNA portion of ribosomes provide only a few hints to the concurrent process of the addition of protein. The step from 43S neosome to 50S ribosome involves both a change in sedimentation coefficient and a change in the elution from DEAE, but no additional RNA is added. It seems quite obvious that this change is due to a change in the protein content. Studies using $C^{14}$-leucine to follow the synthesis of the protein are described in Section VII.

## VI. Kinetic Studies of Phenol-Extracted RNA

The kinetics of incorporation of tracers can also be observed in the purified RNA extracted by phenol. After short exposures to $C^{14}$-uracil some radioactivity is associated with the main peaks of 16S and 23S RNA and the rest appears as a component of average sedimentation coefficient of about 8S (McCarthy and Aronson, 1961; Gros *et al.*, 1961). At the earliest times, however, there is always a considerable fraction of the radioactivity of sedimentation coefficient of 16S or greater even though the specific radioactivity of the 8S region is high (Fig. 22).

Evidently such results are not compatible with any simple precursor product relationship.

Phenol-extracted RNA may also be fractionated by chromatography on columns of methylated serum albumin adsorbed on kieselguhr (Mandell and Hershey, 1960). A clear separation is obtained between soluble RNA, DNA, and high molecular weight RNA. Analyses of pulse-labeled RNA show two main features. At early times almost all of the radioactivity appears in the region of high molecular weight RNA in three peaks, none of which corresponds exactly to the two peaks of ribosomal RNA (Fig. 23). Moreover, the over-all specific radioactivity of the high molecular weight region is greater than that of soluble RNA (e.g., a ratio of 3–4 at 1 minutes) (Fig. 23).

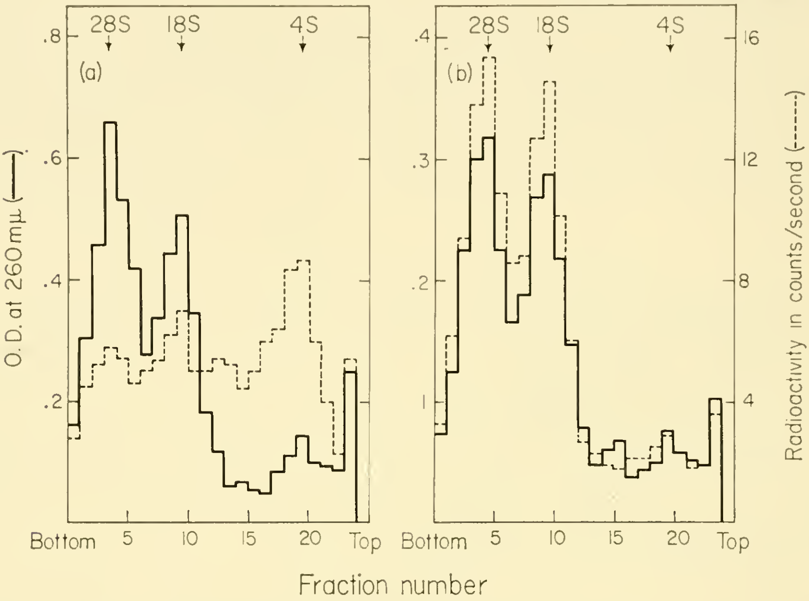


FIG. 22. Sedimentation analysis of RNA extracted from ribosomes (*a*) from cells given a 30-second exposure to $C^{14}$-uracil, (*b*) from cells given a 30-second exposure to $C^{14}$-uracil followed by 20 minutes growth in $C^{12}$-uracil at a 200-fold excess concentration. Centrifugation 230 minutes at 37,000 rpm, 4°C.

Analysis of a number of time points in this way allows determination of the delay of entry of label into S-RNA and DNA relative to the high molecular weight RNA. Parallel experiments carried out with $C^{14}$-uracil and $P^{32}$ as the label lead to a delay time of 1–2 minutes. The interpretation has been made (Midgley, 1962) that the entry of label into both S-RNA and DNA results from nucleotide material produced by the degradation of a part of the high molecular weight RNA fraction. Thus if the D-RNA portion of the eosome fraction were degraded after an average lifetime of some 2 minutes, the delay in labeling could reflect the utilization of these secondhand nucleotides for S-RNA and DNA synthesis.

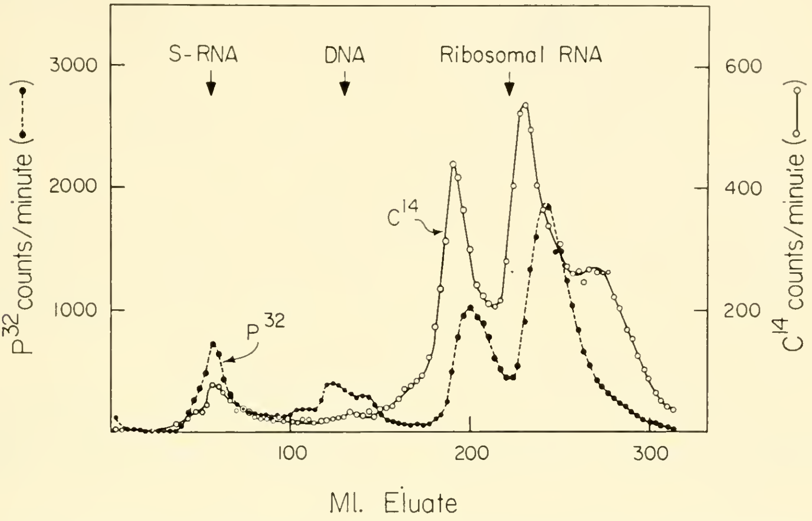


FIG. 23. Chromatography on a column of protein-coated kieselguhr of phenol-prepared nucleic acids from *E. coli* labeled randomly with $P^{35}O_4^{---}$ and for 1 minute with $C^{14}$-uracil. Elution with a linear gradient of NaCl in 0.04 *M* phosphate buffer, pH 6.7, from 0.4 *M* to 1.1 *M*. Flow rate 35 ml/hour.

In support of this view is the fact that the delay in S-RNA and DNA labeling by $C^{14}$-uracil is abolished by the presence of chloramphenicol. Under these conditions, where $C^{14}$-uracil accumulates in the eosome fraction (see Section VII), nucleotide analyses indicate that D-RNA is synthesized but degradation is markedly reduced (Midgley, 1962). There is then little transfer of label from D-RNA to S-RNA and DNA.

## VII. KINETICS OF INCORPORATION INTO PROTEIN

### A. INCORPORATION OF $C^{14}$-LEUCINE INTO THE RIBOSOMES

The separation of ribosomes from their precursors by chromatography depends, presumably, on differences in their proportions of protein. Thus it appears that the product ribosomes are not formed by accretion of material having a constant nucleic acid/protein ratio but that the addition of RNA and protein occurs in time separated stages. More

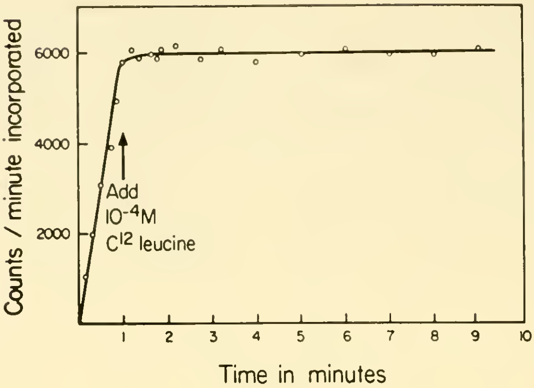


FIG. 24. The incorporation of $C^{14}$-leucine ($10^{-6}$ *M*) into protein in a culture of *E. coli* ML30. $C^{12}$-leucine was added to a final concentration of $10^{-4}$ *M* at 1 minute.

direct evidence for these processes can be obtained by observation of the kinetics of incorporation of amino acids into ribosomes.

These experiments suffer, however, from two technical difficulties. A part of the newly formed protein found associated with ribosomes is not destined to remain in the ribosomes but is nascent protein which is sub-

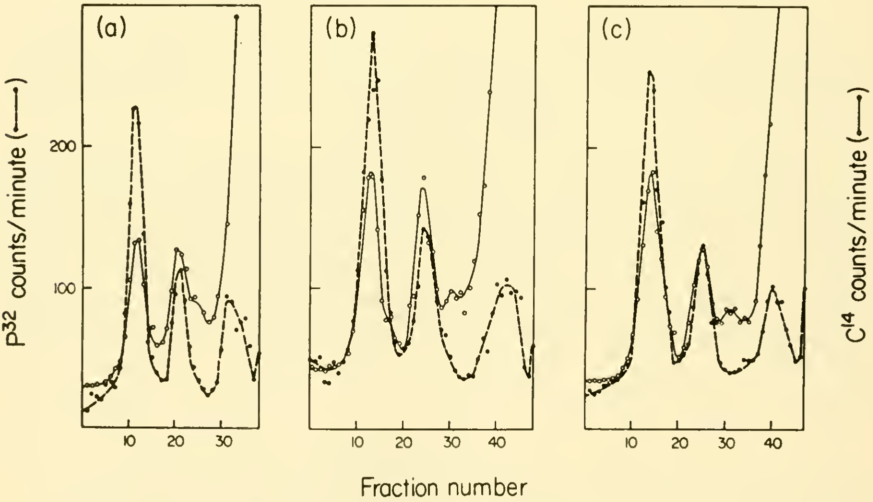


FIG. 25. Sedimentation analysis of three total cell extracts prepared from cells labeled with $P^{32}$ for four generations and $C^{14}$-leucine for 1 minute followed by (*a*) 10 seconds, (*b*) 2 minutes, (*c*) 7 minutes, 45 seconds in $C^{12}$-leucine (see Fig. 24). Cells washed and extracts prepared in tris-HCl 0.01 *M*, pH 7.4, containing $MgCl_2$ $10^{-4}$ *M*. Centrifugation 165 minutes at 37,000 rpm, 4°C. $P^{32}$ shows the location of the 50S and 30S ribosomes.

sequently released as soluble protein (see Section I,E). In addition, that part of the newly formed protein which is destined to be a part of the ribosomes does not appear to be firmly bonded to the RNA because a considerable proportion is not eluted from DEAE. Thus chromatography cannot be used to complement the results of sedimentation analysis and the sedimentation analysis is obscured by the presence of nascent protein.

The time course of incorporation of $C^{14}$-leucine is shown in Fig. 24, taken from Britten *et al.* (1962). Figure 25 shows the sedimentation profiles obtained at various times after a 1-minute period of incorporation was terminated by dilution of the tracer.

The results differ markedly from those obtained with uracil as a tracer. There is no sign of the delay which would be apparent if a considerable pool of precursors to ribosome protein existed. Most of the $C^{14}$-leucine radioactivity appears to enter the 30S and 50S ribosomes directly. In addition, a part of the tracer enters the 43S neosome and is subsequently transferred to the 50S ribosome.

### B. EFFECT OF CHLORAMPHENICOL ON RIBOSOME SYNTHESIS

Chloramphenicol, which inhibits protein synthesis, provides a subsidiary technique for showing which stages of ribosome synthesis depend on the addition of protein. Figure 26 shows that in the presence of chloramphenicol most of the $C^{14}$-uracil tracer is restricted to particles having sedimentation coefficients less than 20S. This RNA has been shown by Bolton (1959) to have ribosomal base composition.

### C. ASSEMBLY OF THE RIBOSOMES

The experiments reported relating to the synthesis of ribosomal protein do not add any fundamentally new features to the sequence of ribosome synthesis. They do, however, confirm other indications that some of the steps in the sequence represent the addition of protein. The fact that a high proportion of the leucine label enters 30S and 50S ribosomes directly with delays of less than 1 minute shows that the last stages of synthesis (30S neosome → 30S ribosome and 43 neosome → 50S ribosome) involve the addition of protein. This is in agreement with the fact that no new RNA is incorporated directly into ribosomes and with the low protein RNA ratio of neosomes as indicated by their column behavior.

One other feature is clear from the leucine pulse and chase experiments. Since the 43S receives some label at early times it is evident that not all the protein of the 50S is added in one step. In fact, the chase experiment shows that the 50S increases some 30% in specific radioactivity during an 8-minute chase period at the expense of the 43S. Rough

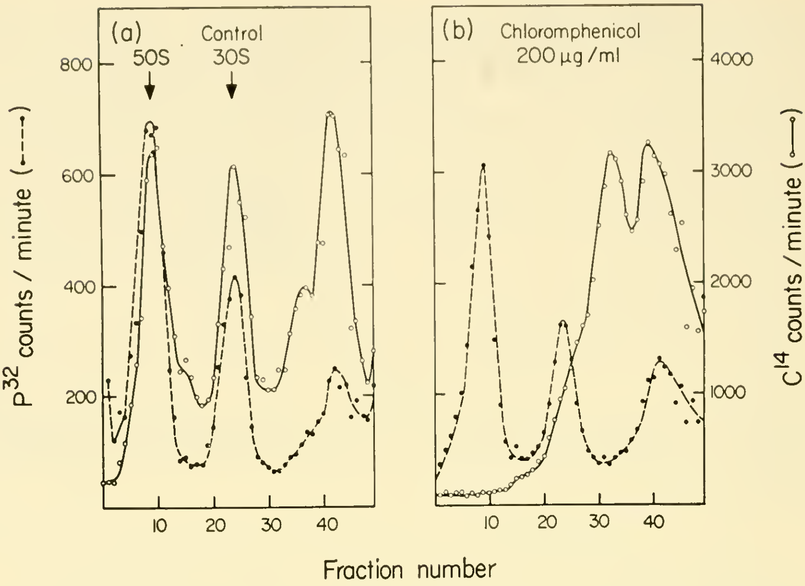


FIG. 26. The effect of chloramphenicol on the incorporation of $C^{14}$-uracil into ribosomes. Cells were grown for three generations in $P^{32}$ and given a 10-minute exposure to $C^{14}$-uracil. (*a*) Control, (*b*) in the presence of chloramphenicol (200 µg/ml added 2 minutes before the uracil). Cells washed and extracts prepared in tris-HCl 0.01 *M*, pH 7.4, $MgCl_2$ $10^{-4}$ *M* in the presence of DNase. Centrifugation 160 minutes at 37,000 rpm, 4°C.

estimates made from the two leucine experiments suggest that the 43S has a protein to RNA ratio ¼ to ⅓ that of the 50S ribosome.

The final flow diagram, including both RNA and protein moieties of ribosomes, is shown in Fig. 27. The open and shaded areas are proportional to RNA and protein contents. The eosome is shown as pure RNA since there are no measurements of its protein content. The 30S neosome is shown with less than half of the protein of the 30S ribosome by analogy to the 43S. This point is not really established because of the lack of resolution between objects in the 30S size range. Likewise, the 30S neosome which is precursor to the 43S is shown as a separate object and this may well be an unnecessary complication. The 43S neosome is shown with only one-quarter of the protein of the 50S ribosome. This quantity is uncertain and it is not known what fraction of its protein enters directly in the formation of the 43S or by way of the 30S precursor to it. While these estimates of relative protein contents in neosome and ribosome are crude and preliminary, it is clear that conversion of neosome to ribosome involves only the addition of protein and that the greater part of the ribosome protein is added in this step.

The effect of chloramphenicol is very much what one would expect from this model. The predominant effect is the accumulation of what appears to be eosome material by both sedimentation and column analysis. Moreover, the base composition of this material is like that of ribosomes. Apparently chloramphenicol inhibits the addition of protein to eosomes. The minor peak appearing at about 24S may represent small quantities of neosome produced in the presence of chloramphenicol since even at 200 $\mu$g/ml the inhibition of protein synthesis is not absolutely complete.

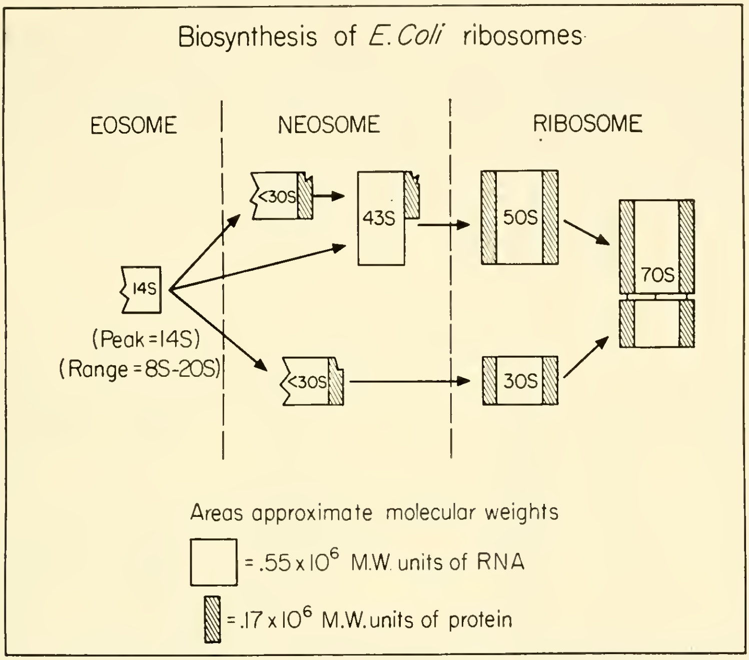

Biosynthesis of E. Coli ribosomes
EOSOME
NEOSOME
RIBOSOME
14S
(Peak = 14S)
(Range = 8S-20S)
<30S
43S
<30S
50S
30S
70S
Areas approximate molecular weights
= .55 x 10^6 M.W. units of RNA
= .17 x 10^6 M.W. units of protein

FIG. 27. The biosynthesis of ribosomes in *E. coli*. The open and shaded areas are proportional to weights of RNA and protein, respectively.

One feature of the general sequence of ribosome synthesis deserves further mention. The addition of RNA and protein occur in time separated stages. This is clear from the very observation of RNA rich intermediates.

## VIII. Composition of Newly Formed RNA

### A. AMBIGUITIES

If an end product is stable, then newly formed material should show the same composition as does the total. If, however, two components of different composition exist and only one of them is stable the composition of newly formed material will differ from the total. The newly formed material must be distinguished by its content of a tracer. Thus the base ratios of newly formed RNA are usually estimated from the $P^{32}$ content of the different nucleotides obtained by hydrolysis of the RNA. This procedure, however, yields an "apparent" composition which may differ from the true composition of newly formed RNA.

If 5′-nucleotides are isolated and measured the apparent composition will be determined partly by the specific radioactivities of the different nucleotide precursors. Several authors have shown that considerable differences in the specific radioactivities of the precursors do exist for prolonged periods of time (Volkin and Astrachan, 1956; Yčas and Vincent, 1960). When 2′,3′-nucleotides are isolated, differences in the specific radioactivities of the precursors will be averaged out provided that the product nucleic acid has a random distribution of nearest neighbors. Accordingly, alkaline hydrolysis to yield 2′,3′-nucleotides is customarily used.

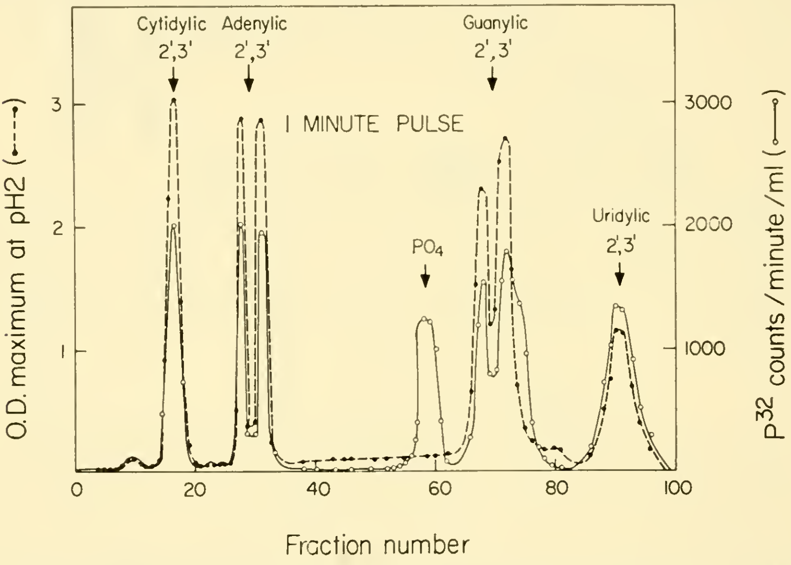


FIG. 28. Separation of nucleotides released by alkaline hydrolysis of RNA by chromatography on Dowex-1. A small quantity of $P^{32}$-labeled unknown material is mixed with a much larger quantity of RNA of known composition before hydrolysis. The composition of the unknown material can be calculated from the measured specific radioactivities of the nucleotide peaks.

The neighbor distribution in ribosomes is random (Bolton, 1960) but appreciable deviations from randomness occur in DNA (Josse *et al.*, 1961). Alkaline hydrolysis should then release nucleotides which show an apparent composition equal to the composition of newly formed RNA if the RNA is similar in its distribution of neighbors to ribosomal RNA. If, however, the RNA is similar to DNA and contains a non-random distribution, the apparent composition will not be an accurate measure of the newly formed material.

At early times after the addition of $P^{32}$ the radioactivity of the free nucleotide pool is much greater than that of the nucleic acid. Some of this radioactivity is likely to contaminate the nucleic acid. It is therefore necessary to use separation techniques capable of separating the 2′,3′-nucleotides released by alkaline hydrolysis from contaminating 5′ nucleotides or serious errors may result. Also, to minimize errors from losses or conversions (by deamination) during hydrolysis it is desirable to measure the radioactivity by isotope dilution (Fig. 28).

### B. TIME COURSE OF COMPOSITION

Differences in the compositions of newly formed material have been observed in several situations. Volkin and Astrachan found the RNA formed after infection by T2 was intermediate in composition between the RNA of the bacteria and the DNA of the phage (Volkin and Astrachan, 1956). The formation of RNA approaching the composition of the DNA was observed in several bacteria (of quite different DNA composition) following a transition from broth to synthetic medium (Hayashi and Spiegelman, 1961).

The same effect also appears in the absence of intentional transients in the growth conditions. Yčas and Vincent observed the apparent composition of the RNA formed by yeast during 5–20 minute periods of exposure to $P^{32}$ (Yčas and Vincent, 1960). Although the data scatter widely, the apparent composition seems to be intermediate between the RNA of the cell and its DNA. Similar results were obtained in a variety of bacteria (Astrachan and Fisher, 1961).

An extensive series of measurements has been carried out by Midgley using several different organisms (Midgley, 1962) (Table VI). The accuracy of these measurements, carried out by isotope dilution, is sufficient to allow an estimate of the time constant of the transition. Correcting for the delays introduced by the nucleotide pool, the time constant is less than 3 minutes, corresponding roughly to that of the eosome stage of ribosome synthesis. Also the composition of the newly formed material (extrapolated to zero time) is not like that of the DNA but is midway between the DNA and the RNA of the cell.

TABLE VI

CHANGE IN APPARENT COMPOSITION OF RNA DEPENDING ON PERIOD OF EXPOSURE TO $P^{32}O_4^{---}$

| Time (min) | A | C | G | U |
|---|---|---|---|---|
| | | *Escherichia coli* | | |
| 0.5 | 25.3 | 23.2 | 29.8 | 21.7 |
| 1 | 24.8 | 23.4 | 30.1 | 21.7 |
| 2 | 25.0 | 22.9 | 29.5 | 22.6 |
| 5 | 26.0 | 22.4 | 29.9 | 21.7 |
| 10 | 25.1 | 22.0 | 30.9 | 22.0 |
| 20 | 25.6 | 20.9 | 32.5 | 21.0 |
| 30 | 25.8 | 21.4 | 33.3 | 19.5 |
| ∞ | 25.2 | 22.1 | 33.5 | 19.2 |
| | | *Bacillus subtilis* | | |
| 0.5 | 25.6 | 23.5 | 27.9 | 23 0 |
| 1 | 25.5 | 22.8 | 27.2 | 24.5 |
| 2 | 25.6 | 23.3 | 27.7 | 23.4 |
| 4 | 26.2 | 23.7 | 28.0 | 22.1 |
| 8 | 26.2 | 22.5 | 29.3 | 22.0 |
| 14 | 25.5 | 22.1 | 31.5 | 20.9 |
| 28 | 25.6 | 22.0 | 31.7 | 20.7 |
| ∞ | 25.5 | 22.1 | 31.4 | 21.0 |

### C. FRACTIONATION

At early times after the addition of the tracer the bulk of the $P^{32}$ is still in the eosome fraction. Thus the composition observed at early times should be a measure of the composition of eosomes. Ribosomes free of S-RNA can be obtained by several cycles of centrifugation in tris buffer containing $10^{-2}$ *M* Mg. The entire eosome fraction can be released from the ribosomes by reducing the magnesium concentration to $10^{-4}$ *M*.

TABLE VII

FRACTIONATION OF NEWLY FORMED RNA[a]

| Released by $10^{-4}$ Mg | Released by $H_2O$ | Not released by $H_2O$ |
|---|---|---|
| C 23.2 | 25.0 | 21.5 |
| A 25.0 | 24.4 | 25.3 |
| G 29.8 | 27.0 | 32.8 |
| U 21.8 | 23.6 | 20.4 |

[a] A pellet of 70S ribosomes was prepared from cells given 2 minutes' exposure to $P^{32}O_4^{---}$. The pellet was then suspended in buffer with a magnesium concentration of $10^{-4}$ *M* or in water with magnesium concentration of $3 \times 10^{-4}$ *M* and centrifuged 40K 240 minutes to remove ribosomes and adhering RNA.

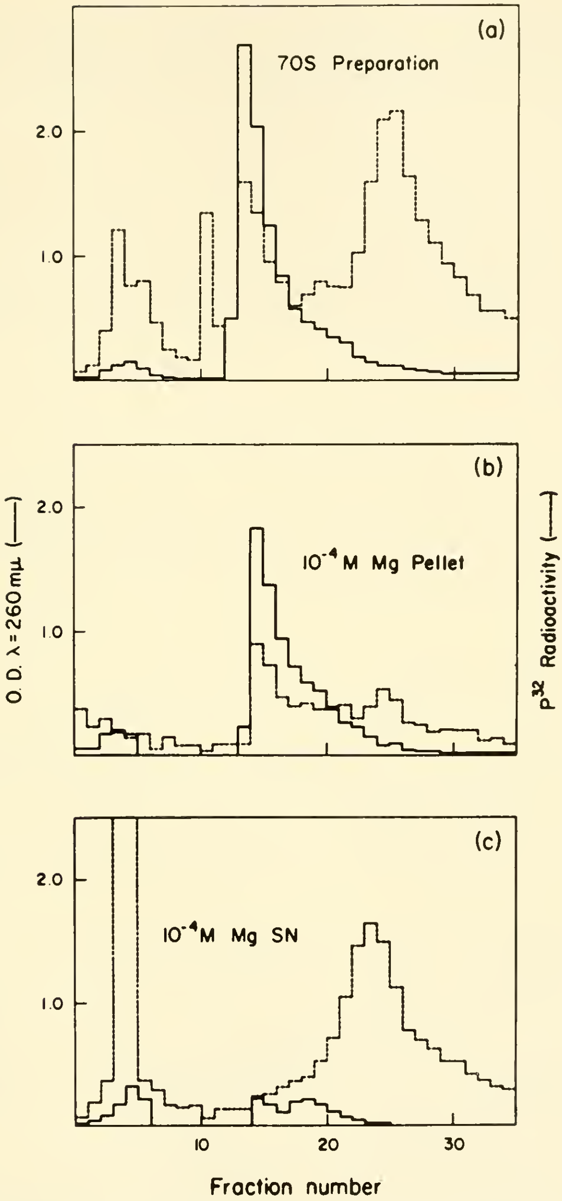


FIG. 29. Purification of the eosome fraction. 70S ribosomes were prepared by centrifugation from *E. coli* given 2 minutes' exposure to $P^{32}O_4^{---}$. A portion was analyzed on DEAE (*a*). Another portion was suspended in buffer with a magnesium concentration of $10^{-4}$ *M* and centrifuged 40K 240 minutes. The pellet (*b*) and supernatant fluid (*c*) were then analyzed on DEAE. Note that the bulk of the material is sedimented but the bulk of the $P^{32}$ remains in the supernatant fluid.

Further centrifugation sediments the ribosomes leaving the eosomes in the supernatant fluid. Chromatography on DEAE provides a final purification of the eosome fraction (Fig. 29). The composition of the material obtained by this procedure is intermediate between DNA and RNA as was expected.

If, however, the ribosome pellet is suspended in water containing a slightly higher concentration of Mg only a portion of the eosome fraction is released. Centrifugation then separates the free eosomes from those still adhering to ribosomes. The compositions of these two fractions differ; the free eosomes are more like DNA and those still adhering to the ribosomes are more like the final RNA product (Table VII) (Midgley, 1961).

### D. RNA CLASSES

The compositions of the 30S and 50S ribosomes and the apparent composition of the newly formed RNA show some systematic relationships. These can be interpreted as indicating that the different types of RNA are the result of different proportions of two hypothetical RNA's

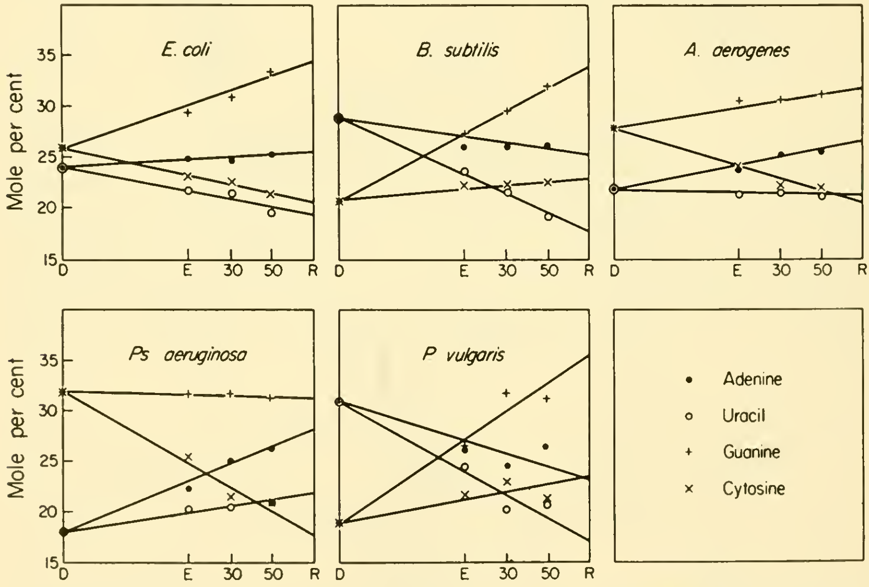


FIG. 30. Relationship between composition of DNA and various RNA fractions. The composition of DNA is plotted at the left (*D*). The compositions of early labeled RNA (E), 30S ribosomal RNA (30), and 50S ribosomal RNA (50) are plotted at the positions expected from the hypothesis described in the text. At the right the lines extrapolate to the composition of a hypothetical purine-rich RNA (R).

which we designate D-RNA and R-RNA to avoid any implication as to their functions. D-RNA indicates an RNA having the same base composition as the DNA and R-RNA indicates a purine rich RNA. For example, the newly formed RNA might be a mixture of equal proportions of D-RNA and R-RNA. The 30S ribosomes having a slightly greater purine excess can then be assumed to contain ⅓ D-RNA and ⅔ R-RNA. The purine excess of the 50S particle is roughly twice that of the 30S ribosome but the kinetics of synthesis indicate that the 30S neosome may be a precursor to the 50S ribosomes. Thus the 50S ribosome may be assumed to contain ⅙ D-RNA and ⅚ R-RNA.

Figure 30 shows that (with the exception of *Proteus vulgaris*) the data are consistent with this hypothesis. The accuracy is not sufficient to rule out other possibilities (e.g., that the 50S contains no D-RNA), but this particular one is the most attractive. The reality of the two types is supported by the experiments showing the fractionation of eosomes. There is, however, no direct evidence that the 30S ribosomes contain a D-RNA component. If this hypothesis happens to be correct, then D-RNA and R-RNA would be formed initially in equal quantities but the requirement for ribosome synthesis would be 2 parts of D-RNA for 7 parts R-RNA.

### E. MECHANISMS FOR THE CHANGE IN APPARENT COMPOSITION

Apparently, more D-RNA is synthesized that could possibly be needed for the production of ribosomes. This interpretation is commonly used as evidence that the D-RNA is subsequently degraded. While this hypothesis may be correct, it poses several problems and it should not be accepted as final until several other alternative interpretations have been excluded.

The degradation products must be such that they are rapidly reutilized without mixing with the large pool of mononucleotides of the cell. If this were not the case, TCA precipitable material would be converted to TCA soluble material when a short period of uracil incorporation is terminated. The expected degradation products (from ribonuclease action) would be 2′,3′-nucleotides which would require phosphorylation to 5′-nucleoside triphosphates before reincorporation. Such a process would be likely to cause mixing with the TCA soluble pool. Perhaps the degradation products are larger and are reutilized without breakdown to mononucleotides by a presently unknown process. Such products could serve as precursors for soluble RNA if the delay of entry into S-RNA is to be attributed to the pool of D-RNA.

As these difficulties are not trivial, it is worthwhile examining some of the alternative interpretations. The concept of degradation and reutilization requires the introduction of the concept of a "private pool," i.e., a pool of nucleic acid precursors separate from the usually observed pool of TCA soluble material. Such a "private pool" could in itself be the cause of the apparent change in composition. If exogenous $P^{32}O_4^{---}$ had a pathway giving direct entry into D-RNA while entry into S-RNA and R-RNA was delayed by a precursor pool, the newly formed RNA would show the composition of D-RNA.

As a third alternative D-RNA could be converted in part to R-RNA by the addition of purines without extensive degradation. Such a process would also result in a change with time of the apparent composition.

At present, no certain choice can be made among these and other possible alternatives. None are completely satisfying and none are completely excluded by the experimental data. The only certainty is that the reactions are complex and elusive.

## IX. DISCUSSION

### A. ROLE OF RIBOSOMES IN INFORMATION TRANSFER

The preceding sections have presented facts concerning the ribosomes, and the flow of material through three stages of ribosome synthesis. These facts provide no direct indications of the part ribosomes may play in directing the formation of protein in accord with the information carried by DNA. However, they provide necessary conditions which must be met by any theory of intermediate processes in information transfer.

At the present time it is easier to provide a reasonable interpretation of a complicated set of biological facts than to demonstrate that one particular interpretation does in fact correspond to nature. For example, it is currently accepted that ribosomes provide the active sites and templates for protein synthesis. The basis for this belief is (*1*) that the rate of protein synthesis is proportional to the ribosome content, (*2*) that ribosomes are essential for protein synthesis in cell free systems, and (*3*) that newly formed protein is associated with ribosomes.

However, one who had reason to believe that ribosomes were not the protein synthesizing sites might interpret the same facts as follows. (*1*) That a high rate of protein synthesis is required for the formation and accumulation of ribosomes, (*2*) that an enzyme required for protein synthesis is firmly bound to ribosomes, and (*3*) that newly formed protein is readily adsorbed by ribosomes.

The correlation of facts relating to the flow of material with another set of facts relating to the flow of information is extremely difficult. In most experiments only the average flow of material into all proteins or all ribosomes can be measured. Furthermore, there may be no immediate alteration in the flow if some of the products are faulty and nonfunctional. In contrast, the experiments dealing with information usually measure one particular function such as an enzymic activity and the rate of synthesis of the particular functional enzyme may differ markedly from the average rate of protein synthesis.

### B. CORRELATION WITH ENZYME INDUCTION

Upon addition of the inducer the rate of synthesis of the enzyme $\beta$-galactosidase increases 200-fold. The time constant of the induction process is as low as 2½ minutes when the inducer level is high. As the same time constant appears in the kinetics of "deinduction" when the inducer is diluted out it seems reasonable to attribute the kinetic characteristics to the formation and subsequent decay of unstable enzyme forming units (Pardee and Prestidge, 1961; Bcezi and Cowie, 1961).

Attributing all of the increase in synthetic rate to an increase in the number of enzyme-forming units (EFU), and assuming that there is 1 EFU per cell before induction, the number of EFU's is about 200/cell in fully induced cells. This number is not unreasonable since $\beta$-galactosidase constitutes 2% of the protein and its synthesis might be expected to involve 2% of the cells' 15,000 70S ribosomes. As the time constant for induction is 2½ minutes, the initial rate of synthesis must be 1.3/sec/cell in order to build up a complement of 200/cell in this time. Since the rate of 70S ribosome synthesis is 3/sec/cell (15,000/cell, 0.02%/sec) the cell could supply completely new ribosomes at the required rate but only by devoting nearly half its capacity for ribosome synthesis to that particular product. Furthermore, if the *de novo* synthesis of complete 70S particles were required, the kinetics of induction might be expected to show the somewhat longer delay characteristic of the appearance of new material in finished ribosomes.

On the other hand, one RNA unit of the eosome fraction might be sufficient to act as template. The minimum number of eosomes synthesized is 36/sec/cell (MW = about $0.15 \times 10^6$, degradation ignored) and the number needed for $\beta$-galactosidase EFU's would be 3.6% of the total. Furthermore, the kinetics of formation of $\beta$-galactosidase EFU's correspond to the kinetic of eosome formation.

The average molecular weight considered is of course too small to specify a protein molecule of 100,000 MW. Perhaps the RNA used as template for $\beta$-galactosidase is considerably larger than the average, or perhaps only a part of the enzyme molecule is dependent upon induction for its synthesis.

In any event the part of the DNA associated with $\beta$-galactosidase must be highly active compared to the average DNA. The entire lac region only comprises a tiny part of the chromosome, yet its product must account for at least 3.6% of the gene-directed material.

This example shows the present difficulties in correlating the facts of ribosome synthesis with their probable role as enzyme-forming units. Only relative enzyme-forming capacities are measured and several wild assumptions (which can only be justified as the best guess at present) are needed to estimate the absolute number of EFU per cell.

### C. MESSENGER THEORY

At present the most precisely formulated theory of the role of ribosomes in information transfer is the messenger hypothesis (Jacob and Monod, 1961). The kinetics of enzyme induction and deinduction indicate that enzyme-forming units are unstable and that template material must be continuously supplied by the gene. However, the ribosomes which appear to be the sites of protein synthesis are stable. Accordingly, Jacob and Monod postulate that an unstable messenger RNA is produced by the gene and associates temporarily with ribosomes to act as template for protein synthesis.

They predict that cells should therefore contain an RNA fraction having the following properties: (*a*) average molecular weight $5 \times 10^5$ or more, (*b*) base composition reflecting the base composition of DNA, (*c*) capable of associating with ribosomes, and (*d*) having a high rate of turnover.

As described in Sections V and VII, material having these characteristics has been found both in experiments designed to identify "messenger" RNA (Brenner *et al.*, 1961; Gros *et al.*, 1961) and in experiments designed to observe ribosome synthesis (McCarthy *et al.*, 1962). Depending upon the objective of the experiment the same material is designated messenger RNA or the eosome fraction. The quantity of this material is between 1 and 3% of the ribosomal RNA. If its molecular weight is $5 \times 10^5$ (⅓ the molecular weight of the RNA of the 70S!) it would be sufficient to provide 3–9% of the 70S ribosomes with templates; only a small fraction of the ribosomes could be active in protein synthesis at any one time. There remains some question whether the "turnover" is due to degradation to low molecular weight material or to incorporation into ribosomes.

The experimental evidence considered in formulating the messenger theory was heavily weighted in favor of the rapid transients observed during enzyme induction or phage infection. Quite a different prediction of the properties of template material could be drawn. Since the amino acid composition of a series of organisms varies much less than does the composition of their DNA's, the composition of the template RNA might be expected to be relatively constant and, therefore, in some cells at least, unlike the DNA. Furthermore, different parts of the DNA may transmit information (i.e., direct the synthesis of RNA) at different rates. The activity of the gene determining the structure of an inducible enzyme appears to be markedly increased upon induction. In this event the composition of DNA-directed RNA should reflect the composition of the DNA weighted in proportion to the activity of the different parts.

Alternatively the whole DNA complement might be replicated as RNA, but subsequently those regions not suitable for protein forming templates would be degraded. In some instances, such as hemoglobin synthesis, the enzyme-forming units are quite stable. Templates might then be expected to show a range of stabilities. Finally, for economical operation of the cell, there should be a sufficient quantity of template material to allow all ribosomes to be active. Thus, as an alternative to the messenger theory, template material might be predicted to have the following properties: (*a*) a constant composition from cell to cell reflecting the average amino acid composition of the proteins, (*b*) having a range of lifetimes depending on the stability of the particular enzyme forming unit, and (*c*) a transient or permanent association with the ribosomes. Material of these characteristics is well known. Ribosomal RNA from bacteria has a relatively constant composition. A wide range of stabilities of template materials can be envisaged if some templates can function only during the eosome stage while others can operate after incorporation into ribosomes.

The value of this model is that it demonstrates clearly that the template material has not yet been identified. The question is still open and cannot be resolved by kinetic studies. At best, the kinetic and compositional studies identify material having the properties *postulated* for "the messenger." Even hybrid formation is not diagnostic of template material as it merely indicates a specific relationship to the DNA. Quite a different approach is needed, such as an assay system to detect template RNA by its effectiveness in synthesizing a definite protein. Nirenberg and Matthaei have made a spectacular beginning in this direction. In their cell-free system polyuridylic acid serves as template for polyphenylalanine and polycytidylic acid for polyproline (Nirenberg and Matthaei, 1961).

### D. HOMOGENEITY OF RIBOSOMES

According to the messenger theory ribosomes could be completely homogeneous except for the attached template component. On the other hand, there might be a specific type of ribosome for each protein of the cell. There are a few indications of some heterogeneity in the ribosomes. The number of different protein components suggests that ribosomes of different protein complement exist (Section II,C). If the ribosomes contain a component approximating DNA in composition, many different types of ribosome are needed to contain even one complete copy of the cell's DNA.

More direct evidence of heterogeneity comes from chromatography of ribosomes. Samples taken from different parts of the elution pattern have different chromatographic properties. These differences may be due to the protein portion of the particles as there is little spread in the elution pattern of the RNA obtained from ribosomes. Slight differences between the RNA's from 30S and 50S ribosomes have been observed both by chromatography and density separations. These differences indicate that both techniques are sensitive to the composition of the RNA and that RNA's of widely different compositions would be detected easily if they had been present.

The most promising technique of isolating a specific class of ribosomes is precipitation by an antibody. At present the only analysis of such RNA shows a marginal difference from the average ribosomal RNA.

Templates for different proteins would not be expected to vary appreciably in their base compositions as the amino acid distribution is relatively constant. An RNA forming the structure of the ribosomes might also be expected to be constant. Thus any separation of ribosomes based on differences in their RNA composition may be difficult to achieve and the heterogeneity of the RNA of ribosomes separated by other techniques (such as antibody precipitation) may be difficult to demonstrate.

### E. SITE OF RIBOSOME SYNTHESIS

Mammalian cells are large enough for radioautographs to show that newly synthesized RNA is localized in the nucleus. Bacteria are too small to allow such direct methods, but Caro and Forro have analyzed the distribution of radioactivity in slices of bacteria. These studies have shown that newly synthesized RNA, unlike the bulk of the RNA, is localized in a central region of the cell (Caro and Forro, 1961). At corresponding times the radioactivity is found in the eosome fraction (see Section IV,B).

The composition of the newly formed RNA suggests that a considerable fraction has a composition like that of the cell's DNA. An enzyme has been isolated which in the presence of a DNA primer catalyzes the polymerization of ribonucleotides into a replica of the DNA (Hurwitz *et al.*, 1960; Weiss and Nakamoto, 1961). Finally, the specific radioactivity of the RNA found in the first pellet (which contains DNA as well as membranes, walls, and unbroken cells) is often unusually high. Thus there are numerous indications that some of the newly formed RNA is synthesized in association with and under the direction of DNA.

### F. DNA-RNA HYBRIDS

Another indication of the direct participation of DNA in the synthesis of RNA is provided by the DNA-RNA hybrids (Hall and Spiegelman; Spiegelman *et al.*, 1961). These complexes form when RNA and the corresponding DNA are heated and slowly cooled. No hybrids are formed unless homologous DNA and RNA are used, demonstrating that complementarity in sequence and not just similarity in composition is required. Such a condition does not seem possible unless the RNA is formed as a direct copy of the DNA.

### G. SELF-REPLICATION OF RIBOSOMES

By far the greater part of the ribosomal RNA has no obvious relationship in its composition to that of the DNA. The composition of the ribosomal RNA is relatively constant in strains showing a wide variation in the GC/AT ratio of DNA (see Section II,C). It is difficult to visualize how RNA originally DNA-like in composition could be so altered to produce a constant purine-rich composition. It is possible, however, that a limited region of the chromosome, which is similar in many bacteria is particularly active in directing the synthesis of RNA. Such RNA might also be stabilized by the prompt addition of basic proteins. If this were the case a single process could be responsible for the synthesis of both D-RNA and R-RNA. Both types would be copies of DNA and the differences would be attributed to rates of synthesis and stability.

The alternative view, that the bulk of the ribosomal RNA is formed by a different process, seems less attractive at present. An enzyme has been found which carries out the appropriate reaction, the polymerization of ribonucleotides under the direction of an RNA primer (Reddi, 1961).

When cells recover from a period of magnesium depletion, during which they lose their complement of ribosomes, the ribosomes increase exponentially. This behavior might be taken to mean that, in the absence of other limiting factors, ribosome synthesis is autocatalytic. There are, however, alternative interpretations.

At present, the self-replication of ribosomes remains a possibility. If, however, it can be established that all ribosomes contain a component whose synthesis is directed by DNA, then self-replication would be ruled out.

### H. CELL-FREE SYNTHESIS OF RIBOSOMES

Several of the cell-free systems for protein synthesis have been reported to show a sensitivity to DNase and to require ribonucleoside triphosphates for optimal synthesis (Kameyama and Novelli, 1960; Matthaei and Nirenberg, 1961; Tissières *et al.*, 1960). Presumably in all these systems there is a DNA-dependent synthesis of RNA which allows a more prolonged period of protein synthesis.

The incorporation of labeled nucleotides into RNA has been observed both in purified enzyme preparations and in some of the cell-free systems capable of protein synthesis. The purified enzyme preparations yield RNA of random sequence (Grunberg-Manago *et al.*, 1955), copies of a DNA primer (Hurwitz *et al.*, 1960; Weiss and Nakamoto, 1961), or copies of an RNA primer (Reddi, 1961), depending on which enzyme is present. As these preparations are not designed to allow protein synthesis, no ribosomes are formed.

In most of the experiments using cell-free systems capable of protein synthesis the RNA concurrently synthesized has not been characterized. In one experiment $P^{32}$-labeled nucleotides were added to the cell free system of Nirenberg and Matthaei. The products (analyzed by chromatography on DEAE) showed RNA eluting at 0.5 $M$ NaCl and at 0.6 $M$ NaCl. These concentrations elute S-RNA and eosomes, respectively (see Section III,A). The peak at 0.6 $M$ NaCl, but not the peak at 0.5 $M$ NaCl, was absent when the system had been treated with DNase. As free RNA elutes at a much higher NaCl concentration, this system seems to be capable of the synthesis of the first ribosome precursor. The composition of the eosome-like RNA showed the typical 50-50 proportions of D-RNA and R-RNA. Thus both D-RNA and R-RNA, but not S-RNA, seem to be synthesized only when DNA is present.

Since these systems have been optimized on the basis of their capacity to synthesize protein it seems likely that ways to improve their capacity for ribosome synthesis could be found. Synthesis of complete ribosomes in cell free systems is a likely development in the future.

### I. STATE OF RIBOSOMES IN LIVING CELLS

It is by no means established that the ribosomes observed in the analytical centrifuge or the electron microscope exist as definite entities of the same form in the living cell. For several years indications have accumulated that ribosomes undergo certain irreversible changes at the instant of breaking the cells.

Whether or not ribosomes appear in electron micrographs of cellular sections depend strongly on the methods of preparation. The pattern observed in the analytical centrifuge is quite different when the cells were broken by lysozyme-freeze-thaw and then put through the pressure cell from the pattern obtained when the processes were interchanged. The quantity of $\beta$-galactosidase associated with the ribosomes can vary 10-fold, depending on the method of breaking. The ribosomes obtained by breaking the cells in the presence of $10^{-2} M$ Mg and dialyzing to $10^{-4} M$ are not identical to ribosomes obtained by breaking in $10^{-4} M$ Mg. The protein content of 30S ribosomes from cells broken in low magnesium concentrations is unexpectedly high. Some find ribonuclease equally distributed between 30S and 50S particles, but others find it only in the 30S particles.

Perhaps all of these anomalies indicate that the ribosome as it exists in the cell has quite a different form and the ribosome as observed in the broken cell juice depends both on the metabolic state of the living cell and on the particular conditions at the time of breakage.

Information can be derived from studies of the particles even though they may be formed at the instant of breakage as they carry reproducible samples of cellular components. In other respects there may be important differences. The low rates of protein synthesis and the difficulties of removing nascent protein from the ribosomes *in vitro* may be due to a less favorable configuration of the particles *in vitro* than *in vivo*. Models of cellular function and structure should not be limited to those conceived in terms of the little round balls observed in the electron microscope. Extended intertwining strands might equally well, or better, represent the organization of ribosomes in the living cell. If this were the case, then at the moment of breakage the strands might gather into their coils quite different proportions of cellular material depending on the conditions at that particular instant.

The freedom given by this point of view is particularly necessary in trying to picture how template RNA might be associated with a preexisting ribosome. The template must contain 600 nucleotides at least if it is to carry information for a polypeptide strand of 200 amino acids. Thus, it is long enough to wrap three times around the 70S particle seen in the electron microscope. It is, however, quite compatible with an extended ribosome.

The extended picture is also more helpful in visualizing ribosome synthesis. The experiments of Caro indicate that newly formed RNA shows the same localization in the cell as does DNA. Thus, eosomes may initially be extended along DNA strands. The hybrid RNA-DNA complexes found by Hall and Spiegelman indicate a considerable binding force may exist. Soon thereafter the eosomes are dislodged, perhaps by accumulation of protein, and can be found associated with ribosomes. Such a transfer would seem more plausible if the recipient ribosome were equally extended.

At present there are no certain facts relating to the state of organization of nucleic acids in the living bacterial cell. A model showing the DNA, the eosomes, neosomes, and ribosomes extended as roughly parallel strands seems equally valid or even preferable to a model in which the DNA is packed in a nucleus surrounded by a multitude of small round ribosomes.

## X. EPILOGUE

In this review of ribosome synthesis we have attempted to fractionate the material and to present the facts in Sections I–VIII partially separated from the interpretations of Section IX. In so doing we have tried to show that many of the currently popular hypotheses are quite acceptable and not in conflict with the facts. They should not, however, be considered as proved until other alternative interpretations have been ruled out. In addition, we have attempted to point out the areas where additional experiments are most needed. We hope that we may have encouraged some of our readers to measure the number of enzyme-forming units, to devise an assay for template RNA, or to carry out a cell-free synthesis of ribosomes.

## REFERENCES

Albertsson, P.-A. (1960). "Partition of Cell Particles and Macromolecules." Wiley, New York.

Aronson, A. I. (1961). *Biochim. et Biophys. Acta* **49**, 98.

Aronson, A. I., and McCarthy, B. J. (1961). *Biophys. J.* **1**, 215.

Astrachan, L., and Fisher, T. N. (1961). *Federation Proc.* **20**, 359.

Boezi, J., and Cowie, D. B. (1961). *Biophys. J.* **1**, 639.

Bolton, E. T. (1958). *Carnegie Inst. of Wash. Yearbook* **57**, 135.

Bolton, E. T. (1959). *Carnegie Inst. of Wash. Yearbook* **58**, 276.

Bolton, E. T. (1960). *Carnegie Inst. of Wash. Yearbook* **59**, 260.

Bolton, E. T., and McCarthy, B. J. (1959). *Carnegie Inst. Wash. Yearbook* **58**, 278.

Bolton, E. T., and Roberts, R. B. (1956). *Carnegie Inst. Wash. Yearbook* **55**, 122.

Bolton, E. T., Hoyer, B. H., and Ritter, D. B. (1958). *In* "Microsomal Particles and Protein Synthesis" (R. B. Roberts, ed.), p. 18. Pergamon Press, New York.

Borsook, H., Deasy, C. L., Haagen-Smit, A. J., Keighly, G., and Lowy, P. H. (1950). *J. Biol. Chem.* **187**, 839.

Brachet, J. (1950). "Chemical Embryology." Interscience, New York.

Brakke, M. K. (1953). *Arch. Biochem. Biophys.* **45**, 275.

Brenner, S., Jacob, F., and Meselson, M. (1961). *Nature* **190**, 576.

Britten, R. J., and McCarthy, B. J. (1959). *Carnegie Inst. Wash. Yearbook* **58**, 267.

Britten, R. J., and McCarthy, B. J. (1962). *Biophys. J.* **2**, 49.

Britten, R. J., and Roberts, R. B. (1960). *Science* 131, 32.

Britten, R. J., Roberts, R. B., and French, E. F. (1955). *Proc. Natl. Acad. Sci. U.S.* **41**, 863.

Britten, R. J., Lark, K. G., and Norcross, F. (1960). *Biochim. et Biophys. Acta* **39**, 150.

Britten, R. J., McCarthy, B. J., and Roberts, R. B. (1962). *Biophys. J.* **2**, 83.

Buchwald, M., and Britten, R. J. (1962). *Biophys. J.*, in press.

Caldwell, P. C., and Hinshelwood, C. N. (1950). *J. Chem. Soc.* p. 1415.

Caro, L. G., and Forro, F., Jr. (1961). *Biophys. Biochem. Cytol.* **9**, 555.

Cassperson, T. (1950). "Cell Growth and Function." Norton, New York.

Chao, F. C. (1957). *Arch. Biochem. Biophys.* **70**, 426.

Chao, F. C., and Schachman, H. K. (1956). *Arch. Biochem. Biophys.* **61**, 220.

Claude, A. (1943). *Science* **97**, 451.

Cowie, D. B., Spiegelman, S., Roberts, R. B., and Duerksen, J. D. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 114.

Dintzis, H. M., Borsook, H., and Vinograd, J. (1958). *In* "Microsomal Particles and Protein Synthesis" (R. B. Roberts, ed.), p. 95. Pergamon Press, New York.

Edelman, I. S., Ts'O, P. O. P., and Vinograd, J. (1960). *Biochim. et Biophys. Acta* **43**, 393.

Elson, D. (1958). *Biochim. et Biophys. Acta* **27**, 216.

Elson, D. (1959). *Biochim. et Biophys. Acta* **36**, 372.

Elson, D. (1961). *Biochim. et Biophys. Acta* **53**, 232.

Elson, D., and Tal, M. (1959). *Biochim. et Biophys. Acta* **36**, 281.

Green, M. H., and Hall, B. D. (1961). *Biophys. J.* **1**, 517.

Gros, F., Hiatt, H., Gilbert, W., Kurland, C. A., Risebrough, R. W., and Watson, J. D. (1961). *Nature* **190**, 581.

Grunberg-Manago, M., Ortiz, P. J., and Ochoa, S. (1955). *Science* **122**, 907.

Hall, B. D., and Spiegelman, S. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 137.

Hall, C. E., and Slayter, H. S. (1959). *J. Mol. Biol.* **1**, 329.

Hanzon, V., Hermodsson, L. M., and Toschi, G. (1959). *J. Ultrastructure Research* **3**, 216.

Hayashi, M., and Spiegelman, S. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 1564.

Hurwitz, J., Bresler, A., and Dinger, R. (1960). *Biochem. Biophys. Research Communs.* **3**, 15.

Huxley, H. E., and Zubay, G. (1960). *J. Mol. Biol.* **2**, 10.

Jacob, F., and Monod, J. (1961). *J. Mol. Biol.* **3**, 318.

Josse, J., Kaiser, A. D., and Kornberg, A. (1961). *J. Biol. Chem.* **236**, 864.

Kameyama, T., and Novelli, G. D. (1960). *Biochem. Biophys. Research Communs.* **2**, 393.

Kurland, C. G. (1960). *J. Mol. Biol.* **2**, 83.

Littlefield, J. W., Keller, E. B., Gross, J., and Zemecnik, P. C. (1955). *J. Biol. Chem.* **217**, 111.

Littlefield, J. W., and Keller, E. B. (1957). *J. Biol. Chem.* **224**, 13.

Luria, S. E., Delbrück, M., and Anderson, T. F. (1943). *J. Bacteriol.* **46**, 57.

McCarthy, B. J. (1959). *Carnegie Inst. Wash. Yearbook* **58**, 281.

McCarthy, B. J. (1960). *Biochim. et Biophys. Acta* **39**, 563.

McCarthy, B. J. (1962). *Biochim. et Biophys. Acta* **55**, 880.

McCarthy, B. J., and Aronson, A. I. (1961). *Biophys. J.* **1**, 227.

McCarthy, B. J., and Britten, R. J. (1962). *Biophys. J.* **2**, 35.

McCarthy, B. J., Britten, R. J., and Roberts, R. B. (1962). *Biophys. J.*, in press.

McQuillen, K., Roberts, R. B., and Britten, R. J. (1959). *Proc. Natl. Acad. Sci. U.S.* **45**, 1437.

Mandell, J. D., and Hershey, A. D. (1960). *Anal. Biochem.* **1**, 66.

Matthaei, J. H., and Nirenberg, M. W. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 1580.

Midgley, J. E. (1961). *Carnegie Inst. Wash. Yearbook* **60**, 305.

Midgley, J. E. (1962). *Biochim. et Biophys. Acta.,* in press.

Midgley, J. E., and Goldthwait, D. (1961). *Carnegie Inst. Wash. Yearbook* **60**, 309.

Nirenberg, M. W., and Matthaei, J. H. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 1588.

Nisman, B. (1961). *Cold Spring Harbor Symposia Quant. Biol.,* in press.

Nomura, M., and Watson, J. D. (1959). *J. Mol. Biol.* **1**, 204.

Nomura, M., Hall, B. D., and Spiegelman, S. (1960). *J. Mol. Biol.* **2**, 306.

Otaka, E., Osawa, S., Oota, Y., Ishihama, A., and Mitsui, N. (1962). *Biochim. et Biophys. Acta,* in press.

Palade, G. E. (1958). *In* "Microsomal Particles and Protein Synthesis" (R. B. Roberts, ed.), p. 36. Pergamon Press, New York.

Pardee, A. B., and Prestidge, L. S. (1961). *Biochim. et Biophys. Acta* **49**, 77.

Pardee, A. B., Paigen, K., and Prestidge, L. S. (1957). *Biochim. et Biophys. Acta* **23**, 162.

Peterman, M. L., and Hamilton, M. G. (1957). *J. Biol. Chem.* **224**, 725.

Reddi, K. K. (1961). *Science* **133**, 1367.

Roberts, R. B. (1958). *In* "Microsomal Particles and Protein Synthesis" (R. B. Roberts, ed.), p. viii. Pergamon Press, New York.

Roberts, R. B. (1959). *Carnegie Inst. Wash. Yearbook* **58**, 290.

Roberts, R. B. (1960). *Ann. N.Y. Acad. Sci.* **88**, 752.

Roberts, R. B., and Duerksen, J. D. (1960). *Carnegie Inst. Wash. Yearbook* **59**, 231.

Roberts, R. B., Britten, R. J., and Bolton, E. T. (1958). *In* "Microsomal Particles and Protein Synthesis," (R. B. Roberts, ed.), p. 84. Pergamon Press, New York.

Schachman, H. K., Pardee, A. B., and Stanier, R. Y. (1952). *Arch. Biochem. Biophys.* **38**, 245.

Schlessinger, D. (1960). *J. Mol. Biol.* **2**, 92.

Spahr, P. F., and Hollingworth, B. R. (1961). *J. Biol. Chem.* **236**, 823.

Spahr, P. F., and Tissières, A. (1959). *J. Mol. Biol.* **1**, 237.

Spiegelman, S., Hall, B. D., and Storck, R. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 1135.

Timasheff, S. N., Witz, J., and Luzzati, V. (1961). *Biophys. J.* **1**, 525.

Tissières, A., Watson, J. D., Schlessinger, D., and Hollingsworth, B. R. (1959). *J. Mol. Biol.* **1**, 221.

Tissières, A., Schlessinger, D., and Gros, F. (1960). *Proc. Natl. Acad. Sci. U.S.* **46**, 1450.

Ts'O, P. O. P., Bonner, J., and Vinograd, J. (1956). *J. Biophys. Biochem. Cytol.* **2**, 451.

Ts'O, P. O. P., Bonner, J., and Vinograd, J. (1958). *Biochim. et Biophys. Acta* **30**, 570.

Volkin, E., and Astrachan, L. (1956). *Virology* **2**, 146.

Waller, J.-P., and Harris, J. I. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 18.

Warren, W., and Goldthwait, D. (1961). *Federation Proc.* **20**, 144.

Weiss, S. B., and Nakamoto, T. (1960). *Proc. Natl. Acad. Sci. U.S.* **46**, 1020; (1961). **47**, 694.

Yčas, M., and Vincent, W. S. (1960). *Proc. Natl. Acad. Sci. U.S.* **46**, 804.

Zubay, G., and Wilkins, M. H. F. (1960). *J. Mol. Biol.* **2**, 105.

## *Chapter* VIII

# Protein Synthesis in Relation to Gene Action[1]

RICHARD SCHWEET AND JOHN BISHOP

*Department of Biochemistry,*
*University of Kentucky College of Medicine,*
*Lexington, Kentucky*

I. Introduction . . . 353
II. Formation of Amino Acyl-RNA Compounds . . . 355
  A. Amino Acid-Activating Enzymes . . . 355
  B. Transfer RNA . . . 356
  C. Specificity of Amino Acyl-RNA Formation . . . 358
III. Steps Requiring Participation of the Ribosome . . . 362
  A. Transfer of Amino Acids from Amino Acyl-RNA to Ribosomes . . . 363
  B. Formation of the Peptide Chain . . . 373
IV. Information Transfer in Protein Synthesis . . . 380
  A. Information Transfer in Bacterial Protein Synthesis . . . 381
  B. Information Transfer in Nucleated Organisms . . . 395
  References . . . 400

## I. INTRODUCTION

The past ten years have seen remarkable advances in understanding the mechanism of synthesis of the major macromolecules of the cell. It appears likely that the major pathways for the synthesis of DNA, RNA, and protein are under study. Some biochemists may consider this viewpoint to be unduly optimistic. Certainly, a vast amount of uncertainty remains. Many new reactions and mechanisms will be discovered. These latter statements are perhaps more true for the field of protein synthesis than DNA and RNA synthesis, despite the great amount of research effort expended in this field. Even for proteins, the main outlines of a major synthetic pathway are apparent. It seems appropriate, therefore, to attempt to summarize the important steps which have laid the foundations for the present state of knowledge.

The comprehensive review by Borsook in 1952 summarized the available information of that date. Among the important earlier advances recognized are the dynamic state of proteins and the energetic requirements for peptide bond synthesis. The use of $C^{14}$-labeled amino acids and *in vitro* (but not cell-free) systems is reviewed and evidence that amino acid incorporation did represent protein synthesis is given in detail. This problem appeared again when cell-free systems were developed. A large part of the review is devoted to analogies between protein synthesis and the biosynthesis of glutamine and glutathione or transpeptidation reactions, since these systems were well studied at the time. The problem of exchange versus *de novo* synthesis also was unresolved at that time as was the status of free peptide intermediates. Both of these problems were raised in part by the "unequal labeling" results of Anfinsen and Steinberg (1951). Some of these points have been clarified; others seem to have lost their significance for the time being.

---

[1] The studies reported here were supported by Grant H-5293 from the United States Public Health Service and Grant G-11400 from the National Science Foundation.

Contributions which turned out to be significant for future developments in protein synthesis were reported just about this time. In Borsook's laboratory, the importance of RNA in protein synthesis was re-emphasized by Holloway and Ripley (1952), and microsomes were shown to be the most rapidly labeled cell fraction (Borsook *et al.*, 1950). Similar results were reported by Hultin (1950). Although a number of experiments on the incorporation of $C^{14}$-amino acids into protein using *homogenates* were reported earlier, the first experiment which clearly implicated microsomes was reported at this time. Siekevitz (1952), in Zamecnik's laboratory, reported that a mixture of mitochondria and microsomes from rat liver incorporated $C^{14}$-alanine at a rate equal to that of the whole homogenate. The role of mitochondria was recognized as providing energy, and the highest amount of incorporation was into the microsomal protein. These findings set the stage for the development of cell-free systems which provided most of the advances in the understanding of protein synthesis in the next ten years.

An assessment of the most significant advances from 1952 through 1961 is more difficult, and reflects undoubtedly some bias on the part of the writers. However, since the details of these studies and particularly the more recent ones will form the subject matter of this review, and some choices must be made to keep the length within reasonable bounds, a summary of these advances will be attempted. The work of Siekevitz (1952) led directly to the system of Zamecnik and Keller (1954), where microsomes plus supernatant fraction plus an ATP-generating system gave good amino acid incorporation. Studies of the function of the supernatant fraction began with the discovery of amino acid-activating enzymes by Hoagland (1955) and by Berg (1956). These enzymes could be assayed in a number of ways, but their physiological role appeared to involve the formation of amino acyl-RNA compounds (Hoagland *et al.*, 1957; Ogata and Nohara, 1957). The existence of amino acyl-RNA compounds was indicated by studies in other laboratories at this same time (see review by Hoagland, 1960). However, the main line of investigation which later led to the conclusion that the amino acyl-RNA compounds were the important soluble intermediate in protein synthesis was initiated by these studies in Zamecnik's laboratory (Hoagland *et al.*, 1957). From these results and those of others came the recognition of a particular type of RNA with a specific function, that is, transfer RNA. Studies of the transfer of amino acids from amino acyl-RNA to ribosomal protein have been reported from a number of laboratories in the past few years and are reviewed in detail below.

The development of two new systems for studying amino acid incorporation led to some of the more recent advances. A cell-free system from rabbit reticulocytes (Schweet *et al.*, 1958b) provided a system where incorporation into a known, soluble protein (hemoglobin) was readily studied. Finally, the bacterial ribosomal preparation (Lamborg and Zamecnik, 1960) made available a system which permitted studies on information transfer in protein synthesis. These studies involving "messenger" RNA are reviewed in detail below.

A number of other pathways for protein synthesis have been reported and some of these are discussed below. However, in this review the transfer RNA to ribosome pathway is emphasized, largely because this is the only system where the intermediate stages have been defined and where the role of nucleic acids has been studied. Thus at least a framework, if not the details, of information transfer from gene to protein can be discussed.

## II. FORMATION OF AMINO ACYL-RNA COMPOUNDS

This subject has been reviewed recently, either as part of general reviews on protein synthesis (Raacke, 1961; Berg, 1961; Campbell, 1961), or in reviews devoted to this topic specifically (Stulberg and Novelli, 1962; Hoagland, 1960). For this reason, only a brief summary of the earlier findings are presented here plus more detailed discussion of recent papers, particularly those relevant to specificity aspects of this reaction.[2]

### A. AMINO ACID-ACTIVATING ENZYMES

These enzymes catalyze the following reactions:

1. Amino acid + ATP + enzyme $\rightleftharpoons$ (amino acyl-Amp-E) + PP
2. (Amino acyl-AMP-E) + T-RNA $\rightleftharpoons$ amino acyl-RNA + enzyme + AMP

3. SUM − amino acid + ATP + T-RNA $\rightleftharpoons$ amino acyl-RNA + AMP + PP

---

[2] The abbreviations are: PP, inorganic pyrophosphate; T-RNA, transfer RNA; A, adenylic acid; C, cytidylic acid; U, uridylic acid; G, guanylic acid; DOC, deoxycholate; C-terminus, free carboxyl terminus; N-terminus, free amino terminus; TMV, tobacco mosaic virus.

Reaction 1 can be measured by the amino acid-dependent incorporation of $P^{32}$-pyrophosphate into ATP and by the formation of amino acyl hydroxamates when large amounts of hydroxylamine are present. The evidence for the formation of the enzyme-bound AMP-amino acid, in which the phosphate is linked in anhydride bond with the amino acid carboxyl, is quite conclusive. This intermediate has been isolated using substrate amounts of enzyme and its synthesis shown by $O^{18}$ transfer experiments.

Reaction 2 has been demonstrated directly starting with the enzyme-AMP-amino acid complex, but is usually measured using the over-all reaction 3. The incorporation of a $C^{14}$-amino acid into T-RNA (transfer RNA, see below) is measured. This reaction is the physiological one catalyzed by these enzymes and they have recently properly been named amino acyl-RNA synthetases (Berg, 1961). The evidence that reactions 1 and 2 are catalyzed by the same enzyme is based largely on the failure to separate these activities on purification of the enzyme (Berg and Ofengand, 1958). Later studies are in agreement with these results, including those with an alanine enzyme where the activity for $PP \rightleftharpoons ATP$ and $AMP \rightleftharpoons ATP$ exchange were purified together (Holley and Goldstein, 1959). The reversibility of reaction 3 has been demonstrated and the equilibrium constant determined from both directions. The values obtained for various amino acids indicate the high-energy nature of the amino acyl-RNA bond (Glassman *et al.*, 1958; Lipmann *et al.*, 1959; Preiss *et al.*, 1959; Leahy *et al.*, 1960). The specificity of amino acyl-RNA synthetases is discussed below. The review by Stulberg and Novelli (1962) contains further details of these reactions.

## B. TRANSFER RNA

The RNA which is active for formation of amino acyl-RNA compounds has been called transfer RNA (T-RNA) based on its function in protein synthesis (Allen *et al.*, 1960) or soluble RNA (S-RNA) based on its location in the cell (Hoagland, 1960). The term T-RNA has been used recently to distinguish this RNA from another type of RNA, also found in the soluble cytoplasm under some conditions, that is, messenger RNA. T-RNA is a specific type of RNA with a molecular weight of about 25,000. However, T-RNA does not appear to be either a precursor or degradation product of larger molecular weight RNA. This conclusion is based on the unique structure of T-RNA. In particular, T-RNA has known end groups as shown:

$$\text{pGp} - - - - - \text{RNA} - - - - - \text{pCpCpA}$$

The discovery of the – – – – – pCpCpA end group was made by a number of laboratories. Enzymes which remove and replace these three nucleotides in a very specific way have been studied (see Hoagland, 1960; Raacke, 1961). There is a rapid turnover of the end group of T-RNA, but the significance of this is not known. The finding of guanosine diphosphate in alkaline hydrolyzates seems to have established this as the major terminus at the other end of the molecule (Singer and Cantoni, 1960; Zillig *et al.*, 1960; Herbert and Canellakis, 1960). In addition, the base composition of T-RNA is unique in its content of "odd" bases. These odd bases, particularly pseudouridine, have been studied by several groups (see Hoagland, 1960 for earlier references). A very complete study of the composition of the soluble RNA of *Escherichia coli* has been reported by Dunn *et al.* (1960). They report the base composition to be (moles per 100 moles): $A = 20.3$, $G = 32.1$, $C = 28.9$, $U = 15.0$, pseudo-U = 2.1, thymine = 1.1, and 2-methyladenine = 0.3. Smaller amounts of other odd bases were also present. Ribosomal RNA contains very little pseudouridine and also differs in having a lower C and higher U and A content. The soluble RNA from other species also contains high proportions of odd bases, but some differences exist. More recently, Cantoni (1962) has reported similar results with rabbit liver soluble RNA, except that thymine nucleotides were not found, but methylamino C was present. Thymine may be an artifact derived from methyl C during hydrolysis (Dunn *et al.*, 1960). The ability of RNA fractions to form amino acyl-RNA compounds was directly proportional to their pseudouridine content (Osawa, 1960). Further discussion of the relation of base composition to function will be given below in connection with the specificity of T-RNA.

The nature of the amino acid to RNA linkage has been studied. This linkage is an ester between the carboxyl group of the amino acid and the ribose hydroxyl of the terminal adenosine. It is not known whether the 2′- or 3′-hydroxyl of the ribose is the esterified group. This finding was first reported by Zachau *et al.* (1958) and was based on the isolation of $C^{14}$-leucyl-adenosine after RNase digestion of $C^{14}$-leucyl-RNA. This result was confirmed and extended by Preiss *et al.* (1959) and Hecht *et al.* (1959). The essentiality of the ACC terminus for amino acid attachment was shown by enzymatic studies and by periodate oxidation of the free *cis*-hydroxyls of the ribose of the terminal adenosine. If an amino acid was attached to the RNA, the ribose was protected from periodate action. This findings led to methods of separating amino acid-specific RNA fractions (see later).

### C. SPECIFICITY OF AMINO ACYL-RNA FORMATION

The specificity of amino acid-activating enzymes for individual amino acids was demonstrated soon after the discovery of these enzymes. A number of enzymes which activate only a single amino acid have been isolated (reviewed by Raacke, 1961; Berg, 1961). Although analogs of the amino acid may be activated by a particular enzyme, in most cases only a single naturally occurring amino acid is activated by a particular enzyme. An exception is the isoleucine enzyme studied by Bergmann *et al.* (1961) which activated valine also when measured by PP-ATP exchange. However, even this enzyme formed only isoleucyl-RNA, a more significant assay for the physiological functioning of these enzymes (Berg *et al.*, 1961). These results thus indicate a two-stage specificity in the action of activating enzymes. First, a specificity for the formation of the enzyme-amino acyl-adenylate complex; and second, a specificity of the complex in its reaction with T-RNA. This second specificity extends to the recognition of T-RNA molecules specific for individual amino acids.

Despite the uniformities in structure at the ends of T-RNA, evidence has accumulated that fractions specific for the attachment of individual amino acids are present. Evidence for T-RNA fractions specific for individual amino acids based on fractionation was first noted by Schweet *et al.* (1958a). Fractionation procedures using untreated T-RNA have resulted in partial separations of fractions specific for various amino acids. Chromatography on an anionic starch resin (Smith *et al.*, 1959), countercurrent distribution (Holley and Merrill, 1959), electrophoresis (Lipmann *et al.*, 1959), and chromatography on calcium phosphate columns of various types (Hartmann and Coy, 1961; Cantoni, 1962) have been used. None of these methods has so far resulted in isolation of a fraction specific for a single amino acid.[3] Promising results have also been obtained by utilizing the change in properties of a particular RNA by virtue of having its amino acid attached. Tyrosyl- and histidyl-RNA chains have been partially purified by complexing to polydiazostyrene columns (Brown *et al.*, 1959). What appears to be a nearly pure RNA fraction specific for valine has been reported by Stephenson and Zamecnik (1961). Yeast soluble RNA containing only attached valine was prepared. After oxidation by periodate of the RNA (the valyl-RNA was protected), and attachment of a bulky dye molecule, fractionation on DEAE-Sephadex yielded a fraction which in one experiment was

---

[3] Holley (private communication) has reported the isolation of amino acid-specific T-RNA fractions using countercurrent distribution.

90% pure. It seems likely that individual fractions of T-RNA specific for single amino acids will eventually be isolated.

A different approach which has provided evidence for amino acid-specific RNA chains was first reported by Hecht *et al.* (1959) and Preiss *et al.* (1959). These authors showed that all T-RNA chains had the same terminal nucleotides (RNA-pCpCpA) and that amino acids were attached to the terminal adenosine. Since individual amino acids were bound to RNA only to a specific extent and there was no competition for sites among various amino acids, amino acid-specific RNA molecules were assumed to be present (based on an unbranched RNA molecule). This approach was amplified by forming a single amino acyl-RNA (valyl-RNA) and then oxidizing with periodate. The amino acid should have protected only the specific RNA. This was tested (Preiss *et al.*, 1959) by removing the bound amino acid with dilute alkali and showing that incorporation of only a single amino acid ($C^{14}$-valine) into the treated RNA was possible. A number of amino acids protected only their own sites against periodate oxidation.

The interest in separation of amino acid-specific T-RNA molecules stems from the hypothesis (Crick, 1958) that the nucleic acid of the ribosome contains a "code" which specifies the amino acid sequence of the protein being made. This code resides in the sequence of the RNA nucleotides. An amino acid, however, according to this idea, cannot be coded for directly by RNA. An "adaptor" molecule which can hydrogen bond with the ribosomal template coding unit is required to transport the amino acid to its proper position in the peptide chain. This adaptor is T-RNA. Each individual T-RNA must therefore carry the code or more likely the complement of the code for the single amino acid it binds. Further discussion of this concept is given below, but its consequences for T-RNA structure and attempts to study these experimentally are appropriate here.

One consequence of the adaptor hypothesis is that differences in amino acid-specific RNA molecules reside in their base sequence and not in secondary structure or molecular weight. This appears to be correct, since T-RNA is fairly homogeneous in size, and while it has a considerable amount of secondary structure, evidence that secondary structure is not required for activity has been reported (see Berg, 1961, for further discussion). Several studies have now appeared which demonstrate considerable variation in the base composition of different T-RNA fractions. An ingenious combination of labeling techniques, enzymatic digestion, and isolation of labeled fragments has been used to study the nucleotide sequences near the adenosine end of T-RNA without prior separation (Berg and Lagerkvist, 1962). The results shown in Table I indicate a remarkable degree of heterogeneity in T-RNA, considering that only the first few nucleotides have been analyzed. These results indicate much greater differences in various T-RNA types than would be expected if the only differences were those in the code, which is expected to be a trinucleotide or in that size range. It is of interest to emphasize that even the molecules which have ---- pCpCpA as their sixth, fifth, and fourth residues (sequences 6 and 10, Table I) are inactive in forming amino acyl-RNA compounds, when the terminal three nucleotides are removed. The complete sequence, in this case ---- pCpCpApCpCpA, is required for activity. Using similar techniques combined with periodate oxidation of chains not protected by attached amino acid, it was also shown by these authors that the chains for isoleucyl-RNA and leucyl-RNA were different. However, the leucyl-RNA contained sequence 2 (Table I) which is present in 25–30% of chains and undoubtedly is not specific for leucyl-RNA. Differences in T-RNA fractions eluted from Brushite columns were found by Cantoni (1962). There was a progressive change in C plus A content and in the "odd" base content of various fractions. Although the purine to pyrimidine ration was 1 in all fractions, the A to U plus pseudo-U ratio departed significantly from unity in some fractions. Thus the resemblance to DNA noted by some authors does not appear in all fractions.

TABLE I
TERMINAL NUCLEOTIDE SEQUENCES OF AMINO ACID-ACCEPTOR RNA[a]

| Sequence[b] | Per cent of RNA chains having sequence shown |
|---|---|
| 1. RNA ---- pGpApCpCpA | 6–8 |
| 2. --------- pGpCpA---- | 25–30 |
| 3. --------- pGpCpU---- | 6 |
| 4. --------- pGpUpA---- | 3 |
| 5. --------- pGpCpCpU---- | 2 |
| 6. --------- pGpCpCpA---- | 4 |
| 7. --------- pGpApG------ | 4 |
| 8. --------- pGpCp(UpC)pA---- | 7–9 |
| 9. --------- pGpCp(UpCpC)p--- | 7–12 |
| 10. --------- pGpApCpCpCpA---- | 4–9 |
| 11. --------- pGpUp(CpCpU)pG---- | 5 |

[a] Taken from Berg and Lagerkvist (1962).

[b] All the sequences terminate in ----pCpCpA as shown in the first line. These have been omitted for simplicity in the lower lines. The exact sequences shown in parentheses are not known.

However, the possibility of a DNA-like structure in localized regions of all T-RNA molecules still remains. From the studies cited above it appears that attempts to study the code by isolating specific T-RNA fractions will be difficult.

Several other complications reinforce this viewpoint. While the evidence points to a particular T-RNA chain being able to accept only a single type of natural amino acid, it appears that more than one kind of chain can code for the same amino acid. This was demonstrated most clearly by the finding that at least 20% of *E. coli* leucyl-RNA had a different base sequence than the rest (Berg and Lagerkvist, 1962). Separation of two fractions from rat liver soluble RNA with leucine-acceptor activity has also been reported (Doctor *et al.*, 1961). Similar results have been reported when T-RNA's from different species have been compared. In general, these experiments have involved studies of the formation of various amino acyl-RNA compounds using T-RNA from one species and crude activating enzyme preparations from another. Thus, different amounts of methionine-RNA were formed using *E. coli* T-RNA depending on whether *E. coli* or yeast enzymes were used (Berg *et al.*, 1961), and guinea pig liver enzymes were quite poor for forming leucyl-RNA with *E. coli* T-RNA, although RNA preparations from a number of mammalian sources were approximately equivalent (Allen *et al.*, 1960). The most detailed study of the species specificity of T-RNA has shown that the amount of a particular amino acid attached to a particular RNA varies greatly, depending on the species from which the activating enzymes are derived and which amino acid is studied (Benzer and Weisblum, 1961). These results with crude enzyme systems are less certain than those involving separation of RNA types. However, it is clear that differences do exist among RNA chains specific for a single amino acid.

These differences may result from a non-universal code. That is, the "coding" base sequence for leucine in T-RNA from rabbits and *E. coli* may be different. A more likely possibility is that the specificity measured in the "species" experiments is for the reaction of T-RNA with the specific amino acyl-AMP-enzyme complex. Species specificity in this case might be expected. The discussion above indicates that a large part, if not the complete T-RNA molecule, is required for the formation of amino acyl-RNA. A smaller part of the T-RNA molecule is expected to form the coding unit (see below), and this may not be the only factor in the reaction with activating enzymes. This explanation for heterogeneity in T-RNA chains specific for a single amino acid does not apply in the case where two leucyl-RNA chains are found in the T-RNA of the same cell and both of these appear to react equally well with a single activating enzyme. It is likely that a degenerate code (more than one code word for a single amino acid) is the explanation here. It should be possible to determine whether these RNA chains contain different codes for leucine by study of the transfer system (see below).[4]

The synthesis of T-RNA is a problem which is so far unresolved. If a specific sequence of bases is involved in the amino acid transfer function of T-RNA, this information might be expected to come from DNA. However, in contrast to results with "messenger" RNA, no system for synthesizing T-RNA is known and even the site of synthesis in the cell is uncertain. Herbert and Canellakis (1961) have reported experiments with soluble RNA which has lost ability to form amino acyl-RNA compounds after treatment with snake venom diesterase. The capacity to accept amino acids was restored by incubation with soluble enzymes and all four (in some cases only three) ribonucleoside triphosphates. Resynthesis of *only* the --- CCA end did not explain the restoration of activity. However, whether a specific base sequence was synthesized, or whether large oligonucleotides were joined, is uncertain. Further studies of this system may clarify the problem of T-RNA synthesis.

## III. Steps Requiring Participation of the Ribosome

A great deal of evidence has accumulated which points to the ribosome as the site of protein synthesis. This evidence consists of kinetic studies where intact cells were incubated with a $C^{14}$-amino acid and at various time periods the specific activity of various cell fractions was examined. The ribosomes were the first to be labeled and then reached a steady state of radioactivity consistent with the hypothesis that most soluble proteins were being synthesized in these particles (Raacke, 1961; Hoagland, 1960). When *specific* proteins were examined, again ribosomes were found to contain protein of highest specific activity (Rabinowitz and Olson, 1959; Peters, 1957; Siekevitz and Palade, 1960; Morris and Dickman, 1960). Studies of "nascent" enzymes on bacterial and yeast ribosomes indicate a similar situation in microorganisms (Kihara *et al.*, 1961; Cowie *et al.*, 1961). The labeling of specific proteins in the cell-free ribosomal system under discussion has also been reported and constitutes the strongest evidence for protein synthesis in these systems (Schweet *et al.*, 1958b; Campbell *et al.*, 1960; Ogata *et al.*, 1960; Suzuka and Shimura, 1960; Kameyama and Novelli, 1960).

In this section we shall attempt to review the evidence concerning various steps which lead to the synthesis of the peptide chain in the ribosome followed by the appearance of the completed protein in the nonsedimentable fraction.

---

[4] Weisblum, Benzer, and Holley have now demonstrated this result (private communication).

### A. TRANSFER OF AMINO ACIDS FROM AMINO ACYL-RNA TO RIBOSOMES

#### 1. *Amino Acyl-RNA as an Intermediate in Protein Synthesis*

The transfer of amino acids from amino acyl-RNA compounds to ribosomal protein indicated that these compounds were intermediates in protein synthesis (Hoagland *et al.*, 1958). Evidence for a soluble intermediate in protein synthesis had been reported earlier by Hultin (1956). The isolation of labeled hemoglobin (Bishop *et al.*, 1961; von Ehrenstein and Lipmann, 1961) and albumin (Hirokawa *et al.*, 1961; Campbell, 1961, 1962), starting with $C^{14}$-amino acyl-RNA, provided strong evidence that these compounds could serve as intermediates for the synthesis of defined proteins.

Kinetic studies with *intact* cells supports the role of amino acyl-RNA compounds as major intermediates in protein synthesis. The work of Hoagland *et al.* (1958) showed that, in ascites cells at depressed temperatures, labeled amino acids were initially linked to soluble RNA more rapidly than they were found in peptide linkage either in the microsome fraction or in soluble protein. Lacks and Gros (1959), using intact *E. coli* cells, also showed that the soluble RNA fraction was most rapidly labeled. In addition, the experiments were compatible with the idea that the label was passed on into peptide linkage when the labeled amino acid ($S^{35}$-methionine) was diluted with a fivefold excess of nonradioactive methionine. When chloramphenicol was added to the incubation, greatly reducing the rate of protein synthesis, the rate of equilibration of the RNA-bound amino acid with the free amino acid pool was also reduced, although the same saturation point was ultimately reached, suggesting that transfer of the RNA-bound amino acid to protein was the normal pathway which released free RNA for further formation of RNA-bound amino acid. Lacks and Gros (1959) feel that amino acyl-RNA is an intermediate in protein synthesis, but that amino acids may *also* extend protein by a second pathway. The hypothesis is based on the observation that the initial rate of labeling of the RNA-bound amino acid was one-third as fast as the initial rate of entry of the label into "protein." Roberts (1959) has indicated that the turnover rate of amino acyl-RNA was too small for these compounds to act as the only intermediate in protein synthesis. These results are consistent with amino acyl-RNA compounds acting as intermediates in protein synthesis in the intact cell, but suggest the existence of other pathways which could be of considerable importance. On the other hand, no other pathway for *protein* synthesis has been demonstrated (see discussion of membrane protein synthesis below). Amino acyl-RNA compounds were shown to be obligatory intermediates for hemoglobin synthesis in a cell-free system (Allen and Schweet, 1960).

### 2. *The Transfer Reaction*

The transfer of amino acids from amino acyl-RNA to ribosomes has now been observed in cell-free systems derived from a number of different sources. Many workers have employed the livers of normal (Hoagland *et al.*, 1958; Nathans and Hülsman, 1960; Moldave and Grossi, 1960; Takanami, 1961) or hepatectomized (von der Decken and Hultin, 1960) rats. Systems from rabbit reticulocytes (Bishop *et al.,* 1960a; von Ehrenstein and Lipmann, 1961), from *E. coli* (Nathans and Lipmann, 1961), and from pea seedlings (Webster, 1960) have also been described. The phenomenon, then, appears to be quite general.

*a. Cofactorial Requirements for Transfer.* A requirement for GTP in the transfer process was first pointed out by Hoagland *et al.* (1958) and has since been confirmed in a number of laboratories (Moldave and Grossi, 1960; Nathans and Hülsman, 1960; Bishop *et al.*, 1960a; Nathans and Lipmann, 1960; von Ehrenstein and Lipmann, 1961). In the reticulocyte system, a maximum rate of transfer is observed in the presence of GTP as the sole phosphorylated nucleoside (Schweet *et al.*, 1961). Takanami (1961) has succeeded in purifying a liver system to the extent that ATP is inactive when it is the sole phosphorylated nucleoside present, while GTP alone is active. Some workers (Hoagland *et al.*, 1958; Grossi and Moldave, 1960) have observed a requirement for ATP in addition to GTP. Magnesium ion is reported to be required in the pea seedling system (Webster, 1960), in rat liver systems (von der Decken and Hultin, 1960; Takanami, 1961), and at a high level in the *E. coli* system (Nathans and Lipmann, 1961).

A requirement for glutathione, or other sulfhydryl compounds, has been demonstrated (Nathans and Hülsman, 1960; Hülsman and Lipmann, 1960) in a rat liver system. In pea seedlings, the glutathione requirement was absolute, and irreplaceable by other sulfhydryl compounds (Webster, 1960); but in *E. coli* the glutathione requirement was less marked than in rat liver (Nathans and Lipmann, 1961). Again in a rat liver system, however, Hirokawa *et al.* (1961) found little requirement for glutathione, presumably because they used an undialyzed pH 5 supernatant (see Hülsman and Lipmann, 1960). More difficult to understand is the inhibitory effect of glutathione observed by Takanami (1961) in a purified rat liver system. However, the purity of commercial glutathione preparations is not always good, and Hülsman and Lipmann (1960) have demonstrated the inhibitory effect of added oxidized glutathione. A strong glutathione requirement was observed in the rabbit reticulocyte system (Bishop and Schweet, 1961a) and shown to be due to the reversible inactivation of an enzyme fraction in the absence of sulfhydryl compounds (Bishop and Schweet, 1961b). A fully activated enzyme, freed of glutathione, could be maintained in the activated state by the addition of Versene. Under these conditions, transfer occurred in the absence of glutathione.

In summary, there is general agreement on a GTP requirement in the transfer reaction, although in two isolated cases ATP was required in addition to GTP for maximum transfer. Whenever it has been studied, a magnesium ion requirement has been found, although the optimum varied from system to system. It is not known whether magnesium ion is required in the transfer reaction itself, or to maintain integrity of the ribosome. Glutathione was required in most cases; in one instance, the requirement was absolute, and in one case glutathione was inhibitory.

*b. Enzymatic Requirements for Transfer.* An enzymatic requirement for the transfer process in the liver system was reported by Hoagland *et al.* (1958) to be located in the pH 5 supernatant fraction. Hülsman and Lipmann (1960), however, have shown this to be almost completely replaceable by glutathione and other sulfhydryl compounds. Subsequently, it appeared that a soluble enzyme fraction was, in fact, required for transfer in the liver system, but that it was present in the microsomes as they were then prepared (Nathans and Lipmann, 1960). By treating their microsome preparation with 0.5% deoxycholate (and thus preparing *ribosomes*) these authors were able to free their preparations of contaminating soluble enzyme and show an almost absolute requirement for added enzyme, even in the presence of glutathione. Rabbit reticulocyte ribosomes, when prepared according to conventional methods, without resort to detergent treatment, are not associated with a lipoprotein membrane. These ribosomes did show a stimulation of amino acid transfer by added enzyme even in the presence of glutathione (Bishop *et al.*, 1960a). However, rabbit reticulocyte ribosomes show an almost *absolute* enzyme requirement following treatment with deoxycholate (Bishop and Schweet, 1961b; von Ehrenstein and Lipmann, 1961). In the pea seedling system (Webster, 1961) and in the *E. coli* system (Nathans and Lipmann, 1961), an enzyme requirement has been observed with washed ribosomes, in the absence of detergent treatment. Thus, it would appear that the enzyme requirement is not an artifact of detergent treatment. Other treatments employed to show an enzyme requirement in rat liver microsomes have been a combination of two detergents (deoxycholate and Lubrol W) and a high KCl level (von der Decken and Hultin, 1960). The same treatment followed by precipitation of the ribosomes with magnesium and dialysis has been used by Takanami (1961).

Using such purified particulate preparations, several assay systems have been developed, and some progress has been made in the purification of the nondialyzable factors required for the transfer reaction. Solubility properties, column chromatography, and heat lability indicate that the transfer factor or factors are protein and presumably an enzymatic reaction is involved (Nathans and Lipmann, 1960; Bishop and Schweet, 1961b). A heat-stable, nondialyzable transfer factor has also been reported (Grossi and Moldave, 1960). Nathans and Lipmann (1960) reported a 30-fold purification of an enzyme fraction from the rat liver 105,000 *g* supernatant by ammonium sulfate and acetone fractionation. Takanami (1961) has purified 8-fold an enzyme fraction from the same source by DEAE-cellulose chromatography. Transfer enzymes have also been purified from pea seedlings (Webster, 1961), from *E. coli* (Nathans and Lipmann, 1961), and from rabbit reticulocytes (Bishop and Schweet, 1961c).

It may be asked whether the transfer function is served by the same group of enzymes which catalyze the activation of the amino acids. This is an extremely difficult question to answer rigorously. It may be that quite different *concentrations* of the same group of enzymes are required for the different types of catalytic activity. Thus, the demonstration that a certain enzyme fraction will catalyze the transfer reaction, but not incorporation from free amino acids (Bishop *et al.*, 1960a; Takanami, 1961), may simply mean that *more* enzyme is required to activate than to transfer the amino acids. However, relative purification data (von der Decken and Hultin, 1960) and the failure to form leucyl-RNA by a fraction which transferred from $C^{14}$-leucyl-RNA to protein (Nathans *et al.*, 1962) is highly suggestive that transfer is not a function of the activating enzymes. The final answer, however, may be found in working with the artificial system of Nirenberg *et al.* (1962) in which a mixture of polyuridylic acid and DNase-treated *E. coli* ribosomes will accept *only* the amino acid phenylalanine from amino acyl-RNA.

The first suggestion that more than one enzyme was involved in the transfer reaction came from von der Decken and Hultin (1960), who fractionated the pH 5 supernatant of rat liver by batch treatment with hydroxylapatite, and followed the activity of various eluted fractions in the transfer of valine and tyrosine. In one particular instance a relative separation of valine and tyrosine transfer activities was observed. However, there was a strong correlation between *low* tyrosine transfer and *high* amount of tyrosine activation, suggesting that the low transfer of $C^{14}$-tyrosine might be due to exchange of $C^{12}$-tyrosine with $C^{14}$-tyrosyl-RNA yielding dilution of the specific activity. Such an exchange seems likely because of the long incubation (in the presence of ATP and a generating system) which was used. On the other hand, the conditions of the incubation, with a very high microsome: amino acyl-RNA ratio, were such that transfer of a single amino acid might well have been observed, in the relative absence of transfer of another. Bishop and Schweet (1961c) fractionated the 105,000 *g* supernatant of the rabbit reticulocyte lysate using salting out with ammonium sulfate, adsorption to calcium phosphate gel, and chromatography on DEAE-cellulose; they obtained two enzyme fractions which showed a 6- to 10-fold enhancement of activity when assayed together. These fractions showed small differences in relative activities in the transfer of four amino acids under conditions (3-minute assay in the absence of ATP or an ATP-generating system) which precluded the occurrence of a significant exchange reaction between RNA-bound and free amino acids. Fessenden and Moldave (1961), using rat liver ribosomes, have reported that the enzymes removed from microsomes with deoxycholate did not catalyze transfer unless a purified, heat-stable, but nondialyzable fraction of the soluble cytoplasm was added. The significance of these fractions is not apparent, since microsomes have been reported to contain transfer enzymes and to show full activity if glutathione alone is added (Hülsman and Lipmann, 1960). Thus, the fraction solubilized by DOC should contain all the transfer enzymes.

On the other hand, Takanami (1961), using a rat liver system, and Nathans and Lipmann (1961) with *E. coli*, have fractionated the supernatant factor on DEAE-cellulose without observing any separation of transfer activities. It appears likely that more than one enzyme is involved in the transfer reaction. However, whether separate stages in the reaction are catalyzed by different enzymes, or whether amino acid-specific fractions have been separated is still not clear.

### 3. *The Mechanism of the Transfer Reaction*

Evidence that transfer RNA is repeatedly recharged with amino acids, after discharging bound amino acid on the template, was presented by Allen and Schweet (1960). The complete reticulocyte cell-free system (starting with free amino acid) showed a considerable requirement for added transfer RNA which was fulfilled by 45–50 $\mu$g of RNA. The stimulation of $C^{14}$-leucine incorporation into protein, however, was 15–20 times as much as the amount of leucine required to charge completely 50 $\mu$g of transfer RNA. Thus, the added transfer RNA must have recycled many times in the course of the incubation. Independently, Nathans and Lipmann (1961) have reported that after discharging its amino acid, *E. coli* RNA could be reisolated (by phenol extraction) and completely recharged.

The original "adaptor hypothesis" (Crick, 1958) called for a relatively small polynucleotide as the adaptor. It was suggested (Hoagland, 1960) that the relatively large size of the transfer RNA molecule was a requirement for specific interaction with the activating enzyme and that, after accepting the amino acid, the transfer RNA molecule might release a small polynucleotide fragment with the amino acid attached to it. The results cited above suggest that the transfer RNA does not break down. Hoagland and Comly (1960) have provided careful studies which point to the same conclusion. $C^{14}$-leucyl-RNA was incubated with various enzyme fractions, or microsome extracts, in the presence and absence of GTP, but in the absence of microsomes. At various times parallel assays were made of the remaining acid-precipitable amino acid, and of the competence of the complete incubation mixture in the transfer process. It was found in every case that the amount of $C^{14}$-leucine transferred to peptide linkage was proportional to the amount of acid-precipitable $C^{14}$-leucine remaining. Thus, free, acid-soluble, amino acid polynucleotide compounds do not appear to be significant intermediates in protein synthesis. It remains to be seen whether such intermediates are formed in association with microsomes.

*a. The Transfer of Labeled, Soluble RNA to the Ribonucleoprotein Particles.* If, as discussed above, amino acyl-RNA fails to form a smaller amino acyl polynucleotide, a direct physical association of the transfer RNA and the microsomes, if only a very transient one, must occur during transfer. A temperature- and energy-dependent transfer of labeled, soluble RNA to ribosomes has been reported by von der Decken and Hultin (1958), by Bosch *et al.* (1959), and by Hoagland and Comly (1960). In all cases, a considerable non-temperature-dependent, non-energy-dependent transfer of labeled RNA to the 105,000 *g* pellet was also observed. The system studied in greatest detail is the rat liver system of Bosch *et al.* (1959, 1960; Bloemendal *et al.*, 1960, 1961). These authors found that a $P^{32}$-S-RNA fraction isolated on an Ecteola column was readily incorporated into ribosomal RNA. The fraction was able to accept $C^{14}$-leucine in the presence of an activating system. After incubation with $P^{32}$-S-RNA, the microsomes were sedimented and the microsomal RNA extracted with phenol and chromatographed on an Ecteola column. Three fractions were recovered, only one of which was labeled. Significantly, the labeled fraction had different chromatographic properties from those of the labeled soluble RNA. Furthermore, when microsomal RNA was isolated after $P^{32}$-labeling *in vivo*, radioactivity was found in all three chromatographic fractions. Bloemendal *et al.* (1961) observed a similar transfer of labeled RNA to detergent-treated liver ribosomes. When the soluble RNA was prelabeled in the *terminal* position with $C^{14}$-adenine and incubated with ribosomes, the $C^{14}$-adenine was found in *internal* positions in the ribosomal RNA (63% of the radioactivity was isolated as adenosine monophosphate after alkaline hydrolysis). It appears, then, that the transfer which occurred in this system resulted in the covalent linkage of soluble RNA with microsomal RNA. It is not clear, however, that this reaction is related in any way to the transfer of amino acids from amino acyl-RNA. In fact, the addition of soluble enzymes and an amino acid mixture failed to stimulate the RNA transfer (Bloemendal *et al.*, 1961).

A transfer of labeled S-RNA to ribosomes has been reported by Hoagland and Comly (1960). In this case, $P^{32}$-labeled, soluble RNA was coupled with $C^{14}$-leucine so that the transfer of both the RNA and the amino acid could be studied. It was shown not only that the $P^{32}$-labeled RNA would enter the sedimentable fraction, but also that it could be displaced by unlabeled soluble RNA. A measure of specificity was indicated by the inability of unlabeled microsomal RNA to displace the labeled soluble RNA from the microsomes. In agreement with the experiments of Bloemendal *et al.* (1961), when the $C^{14}$-leucine was removed from the $P^{32}$-RNA, no inhibition of RNA transfer was observed.

These various experiments do demonstrate that an association between soluble RNA and the ribonucleoprotein particles may be very readily achieved. The slight stimulations observed upon the addition of ATP or GTP may well indicate simply that this is not an energy-dependent step.

Since it is altogether likely that a physical association can occur between soluble RNA and the microsomes, it *might* be expected that when $C^{14}$-amino acid-labeled amino acyl-RNA is incubated briefly with microsomes, some portion of the amino acyl-RNA will sediment with the microsomal fraction. Zamecnik (1960) has in fact observed such a transfer to the RNA of the sedimentable fraction in a rat liver system. The problem of the first intermediate in the ribosome has been studied in the reticulocyte system starting with $C^{14}$-leucyl-RNA (G. Favelukes and R. Schweet, unpublished data) to avoid complications of contamination with free $C^{14}$-amino acid. Incorporation into ribosomes in alkali-labile linkage was used as a measure of amino acyl-RNA binding. As shown in Table II, a GTP- and enzyme-dependent linkage was found. However, the significance of these alkali-labile counts as an intermediate in protein synthesis is dubious, since these counts could not be "chased" into alkali-stable (protein) linkage. It seems likely, therefore, that while there must be some initial association of amino acyl-RNA and ribosomes, this is very transitory and the amino acid is transferred from the amino acyl-RNA to an alkali-stable form in the ribosome very rapidly. The RNA is then free to form amino acyl-RNA once again.

TABLE II

RIBOSOMAL $C^{14}$-LEUCINE IN ALKALI-LABILE LINKAGE[a]

| Assay conditions[b] | $C^{14}$-Leucine incorporated Labile ($\mu\mu$moles) | $C^{14}$-Leucine incorporated Stable ($\mu\mu$moles) |
|---|---|---|
| Complete | 93 | 154 |
| Minus enzyme | 4 | 6 |
| Minus GTP | 20 | 9 |
| Complete plus $2 \times 10^{-5}$ *M* puromycin | 74 | 27 |

[a] G. Favelukes and R. Schweet (unpublished data).

[b] DOC-ribosomes (2.8 mg), $C^{14}$-leucyl-RNA (0.3 mg), enzyme (3 mg) and GTP (1 $\mu$mole) were incubated at 37°C for 3 minutes.

*b. The Role of GTP in the Transfer Reaction.* An understanding of the role of GTP is essential to the elucidation of the mechanism of the transfer reaction; but to date little information is available. Nathans *et al.* (1962) have shown, in the reticulocyte transfer system, that the breakdown of GTP to GMP and pyrophosphate is several orders of magnitude lower in extent than amino acid transfer. This would appear to eliminate the breakdown of GTP to GMP as a significant factor in the transfer process. Webster (1961), using a pea seedling system, has reported a reversible breakdown of GTP accompanying amino acid transfer. Incubation of ribosomes charged with amino acids and soluble RNA resulted in the formation of amino acyl-RNA and GTP. On the other hand, Nathans *et al.* (1962) have observed a rapid breakdown of GTP to GDP upon mixing *E. coli* ribosomes and soluble enzymes in the *absence* of amino acyl-RNA.

Several pieces of evidence which suggest a role for GTP have come from studies of "release" in the reticulocyte ribosome system (Morris and Schweet, 1961; Morris *et al.*, 1962). In these studies it was shown that ribosomes prelabeled with $C^{14}$-leucine could be incubated in a complete system with $C^{12}$-amino acids and the radioactivity "chased" into soluble hemoglobin. In this system, incorporation of $C^{12}$-amino acids finishes partially completed chains in the ribosome and the formation of labeled, soluble protein is a measure of these reactions. It has now been found that GTP will also promote release of soluble $C^{14}$-protein from prelabeled ribosomes (Table III). This formation of soluble protein occurs in the absence of incorporation of free $C^{14}$-amino acids into ribosomes. This result demonstrates a role for GTP at a later stage than the transfer of amino acid from amino acyl-RNA. One interpretation of this finding is that GTP acts to convert intraribosomal intermediates into peptide linkage, thus producing completed or almost completed chains which are then released. It is postulated that a GTP is split in the process of forming each peptide bond and that GTP does not act in the initial transfer step from amino acyl-RNA. However, since ribosomes as isolated are filled with intermediates, no transfer from amino acyl-RNA is observed in the absence of GTP, since no ribosomal sites are available. Thus, an intraribosomal reaction requiring GTP is needed in order for an earlier step, the transfer from amino acyl-RNA, to occur. This mechanism is advanced as a working hypothesis to guide future studies of the role of GTP.

TABLE III

REQUIREMENTS FOR INCORPORATION AND RELEASE[a]

| Assay components | Incorporation (%) | Release (%) |
|---|---|---|
| Complete system | 100 | 100 |
| No energy | 1.5 | 4.1 |
| Plus 1 μmole ATP only | 51.1 | 50.2 |
| Plus 1 μmole GTP only | 4.9 | 35.2 |

[a] A. Morris and R. Schweet (unpublished data).

### 4. *Specificity in the Transfer Reaction*

A lack of specificity for the transfer RNA species is the most striking aspect of this reaction. Von Ehrenstein and Lipmann (1961) demonstrated that *E. coli* $C^{14}$-amino acyl-RNA could form hemoglobin when incubated with rabbit reticulocyte ribosomes. Bishop *et al.* (1961) showed that amino acyl-RNA from guinea pig liver was competent in hemoglobin synthesis with reticulocyte ribosomes. In these experiments, the labeled hemoglobin was digested with trypsin and fingerprinted. The correspondence of leucine-labeling in these peptides with whole-cell labeled hemoglobin (Table IV) provides good evidence for the synthesis of hemoglobin from this heterologous amino acyl-RNA.

Despite the ability of guinea pig liver amino acyl-RNA to act as a precursor of hemoglobin synthesis, it was observed (Bishop *et al.*, 1961) that amino acids were transferred more readily from amino acyl-RNA to homologous ribosomes. Guinea pig liver transfer RNA was labeled with $H^3$-leucine, and rabbit reticulocyte RNA with $C^{14}$-leucine. To eliminate possible effects of contaminants of the RNA preparations, the two labeled RNA's were mixed and incubated in parallel experiments with either rabbit reticulocyte or guinea pig liver ribosomes. More $H^3$-leucine than $C^{14}$-leucine was transferred to guinea pig liver ribosomes, and more $C^{14}$-leucine that $H^3$-leucine to reticulocyte ribosomes. Two explanations of these results are suggested. (*1*) There exist subtle differences between rabbit reticulocyte and guinea pig liver leucine-transfer RNA which affect the *rate* of the transfer reaction without influencing the *site* to which the leucine is transferred. (*2*) Different types of transfer RNA, specific for the *same* amino acid, are present and have different *transfer* specificities, perhaps transfer to different types of sites. If, for example, two kinds of leucyl-RNA's exist, but in different relative proportions in guinea pig liver and rabbit reticulocytes, the results may be explained if, in addition, different relative frequencies of two corresponding types of *template* sites are found in the ribosomes of these tissues.

TABLE IV

RADIOACTIVITY OF PEPTIDES OF ALPHA CHAIN OF RABBIT HEMOGLOBIN[a,b]

| | Fingerprint 1 | | Fingerprint 2 | |
|---|---|---|---|---|
| Peptide No. | Intact cells ($C^{14}$-leucine) | Cell-free ($H^3$-leucine) | Intact cells ($C^{14}$-leucine) | Cell-free ($H^3$-leucyl-RNA) |
| 1 | 2.5 | 4.6 | 9.2 | 8.7 |
| 2 | 8.2 | 5.1 | 3.8 | 1.9 |
| 3 | 31.1 | 12.5 | 32.2 | 36.9 |
| 4 | 7.2 | 6.6 | 6.1 | 6.8 |
| 5 | 18.2 | 20.4 | 20.3 | 20.4 |
| 6 + 7 | 1.3 | 2.0 | 1.8 | 1.1 |
| 8 | 3.2 | 2.3 | 1.4 | 1.6 |
| 9 | 0 | 0 | 0 | 0 |
| 10 | 16.4 | 19.8 | 13.5 | 13.0 |
| 12 | 12.2 | 26.5 | 11.6 | 9.7 |
| 13 | 0 | 0 | 0 | 0 |
| 14 | 0 | 0 | 0 | 0 |
| Blank 1 | 0 | 0 | 0 | 0 |
| Blank 2 | 0 | 0 | 0 | 0 |
| Blank 3 | 0 | 0 | 0 | 0 |

[a] From Bishop *et al.* (1961).

[b] Radioactivity is expressed as the percentage of the total radioactivity in all peptides. 0 denotes less than 0.1%.

Specificity in transfer enzymes has been reported. Transfer enzyme fractions of rabbit reticulocytes and *E. coli* were not interchangeable

(Nathans and Lipmann, 1961). The partially purified *E. coli* enzyme fraction already referred to was completely inactive in the transfer of amino acids from amino acyl-RNA to reticulocyte ribosomes. Similarly, rabbit reticulocyte pH 5 supernatant was inactive in the *E. coli* system. The significance of this type of specificity is not understood.

### B. FORMATION OF THE PEPTIDE CHAIN

Understanding of the mechanism of assembly of the amino acids may be expected to yield information on the central problem of how amino acid sequence is specified. One approach has been to study the distribution of a $C^{14}$-amino acid in various portions of a specific protein synthesized in the presence of this labeled amino acid. If the specific activity of the $C^{14}$-amino acid is different in various parts of the peptide chain (unequal labeling), information about intermediate pools or sequence of peptide bond formation can be deduced (Loftfield, 1960).

#### 1. *Labeling Patterns in Intact Cells*

Unequal labeling of specific proteins has been observed in a number of systems. The clearest example, perhaps, is the unequal labeling of the ovalbumin synthesized in chicken oviduct (reviewed by Steinberg *et al.*, 1956). Unequal labeling was clearly demonstrated by analysis of the products of (*a*) partial acid hydrolysis, (*b*) peptic digestion, and (*c*) sequence analysis of the hexapeptide liberated by subtilisin. The data were interpreted to indicate the existence of intermediate pools of some kind, or an ordered synthesis of the chain. It is not possible to distinguish between these possibilities as the author noted. Equal labeling of goat lactalbumin was observed by Askonas *et al.* (1955). This is consistent with any of the hypotheses discussed above, since the time to synthesize a lactalbumin molecule was not determined.

The $\alpha$-amylase secreted by late log phase *B. subtilis* is the subject of a careful study by Yoshida and Tobita (1960). $C^{14}$-leucine was added to the culture during the later stages of growth, $\alpha$-amylase was isolated, and leucine was isolated from positions close to the C-terminus, the N-terminus, and from eight tryptic peptides. The highest specific activity leucine was found close to the C-terminus, and the lowest close to the N-terminus. Furthermore, the specific activities of the leucine in the eight tryptic peptides fell between these limits. This experiment, then, is suggestive of an *ordered* synthesis of the polypeptide chain of $\alpha$-amylase, starting at the N-terminal end, and finishing at the C-terminal end. It should be noted, however, that a high molecular weight precursor of amylase was present before the isotope experiment began and the results may mean that this precursor was completed.

The labeling of hemoglobin in reticulocytes has been studied by a number of workers, with rather conflicting results. Muir *et al.* (1952) injected rats with $C^{14}$-valine and demonstrated that the specific activity of the N-terminal valine of isolated hemoglobin was the same as the specific activity of the valine of the remainder of the molecule. On the other hand, Kruh *et al.* (1957, 1960) using radioactive glycine and phenylalanine have observed unequal labeling of hemoglobin in isolated rabbit reticulocytes and in the living rabbit. At times after the addition or injection of the labeled amino acid ranging from 1 to 4 hours in experiments with isolated cells, and from 5 to 58 days in the living animal, hemoglobin was isolated. The specific activity of the labeled amino acid released after a brief hydrolysis of each sample with HCl was compared with the specific activity after complete hydrolysis. In the experiments with isolated cells, unequal labeling could still be detected after 4 hours, while in the living rabbit uniform labeling was not attained until between 30 and 50 days following injection. These experiments are difficult to understand in terms of any of the hypotheses already put forward. Since the time for synthesis of the hemoglobin molecule is estimated to be approximately 1 minute (Dintzis *et al.*, 1958; Schweet *et al.*, 1961), *ordered* synthesis could not result in significantly unequal labeling at times exceeding 20 minutes. Schapira *et al.* (1960) have examined the question in another way. After incubation of reticulocytes with radioactive arginine, hemoglobin was isolated and digested with trypsin. The digest was "fingerprinted," the arginine peptides identified, eluted, and hydrolyzed. The specific activity of the arginine in each of nine peptides was found to be different. It should be noted that Jonxis (1958) has reported four arginine residues per half-molecule of hemoglobin. It appears unlikely that the long-term unequal labeling was due to an analytical artifact since unequal labeling was detected in two different ways. It is possible that slow, long-term exchange reactions, or the presence of several types of hemoglobin with different turnover rates, could explain the data. Further work is needed to clarify these results.

### 2. *Evidence for Ordered Peptide Bond Formation*

As mentioned earlier, labeled amino acids rapidly appear in ribosomes in alkali-stable linkage, presumably protein (Hoagland, 1960). In intact reticulocytes, "pulse-labeling" experiments demonstrated the transitory nature of this material and indicated that hemoglobin precursors, rather than the structural protein of ribosomes, were formed (Dintzis *et al.*, 1958; Rabinowitz and Olson, 1959). These results suggested the presence of peptides in the ribosome which were hemoglobin precursors. Similar results were obtained in the cell-free system from reticulocytes (Schweet *et al.*, 1958b).

Since earlier studies had shown that the labeled ribosomal material could be "chased" directly into soluble hemoglobin when labeled ribosomes were incubated in the complete cell-free system (Schweet *et al.*, 1961; Morris and Schweet, 1961), this system was used to study the nature of the ribosomal intermediates. Ribosomes were labeled in the intact reticulocyte with $C^{14}$-valine, the N-terminal amino acid of hemoglobin. These ribosomes were washed and then incubated with $C^{12}$-valine in the cell-free system. Comparison of the amount of $C^{14}$-valine in N-terminal positions and in internal positions in the soluble hemoglobin formed, indicated that the ribosomes contained incomplete chains synthesized sequentially from the N-terminal end (Bishop *et al.*, 1960b). This was also demonstrated in intact reticulocytes by similar methods (Schweet *et al.*, 1961).

Dintzis (1961) studied the labeling in peptides formed by the tryptic digestion of the separated chains of hemoglobin labeled in the intact cell. He showed that tryptic peptides isolated from hemoglobin after short periods of incubation with labeled leucine could be arranged in a sequence of increasing specific activity relative to the specific activities of the same peptides from hemoglobin isolated after 6 hours of incubation. The peptide with the highest specific activity after short periods of labeling was near the C-terminus. These results are also suggestive of synthesis from the N-terminus is sequential fashion, although other interpretations of labeling data using intact cells are possible (see Steinberg *et al.*, 1956). The results of Goldstein and Brown (1961) indicate that this mechanism may be general. A leucine-requiring strain of *E. coli* was starved of leucine to reduce the possibility of interference by unlabeled pools and then labeled for very short times with $C^{14}$-leucine. The shorter the labeling period, the more $C^{14}$-leucine was found near the C-terminus. On the other hand, Shimura *et al.* (1956) have shown that shortly after injection of $C^{14}$-glycine into silkworm larvae, the N-terminal glycine of silk fibroin had the highest specific activity. A sequential synthesis of many proteins from the N-terminal end appears likely.

The significance of polypeptide chain synthesis from the N-terminal end, if any, is not known; however, if one makes several reasonable assumptions some interesting consequences follow. The two main assumptions are:

1. In addition to functioning to position amino acids in the proper sequence for the protein being made, the ribosome functions to position amino acids in the proper *steric* relation so that the peptide bond can be

formed between the amino group of one amino acid and the carboxyl group of the other. The proper condition for this is shown in Fig. 1A. Two carboxyl-activated amino acids are attached to the ribosome ready for peptide bond formation. These may be attached to transfer RNA, or in some other linkage.

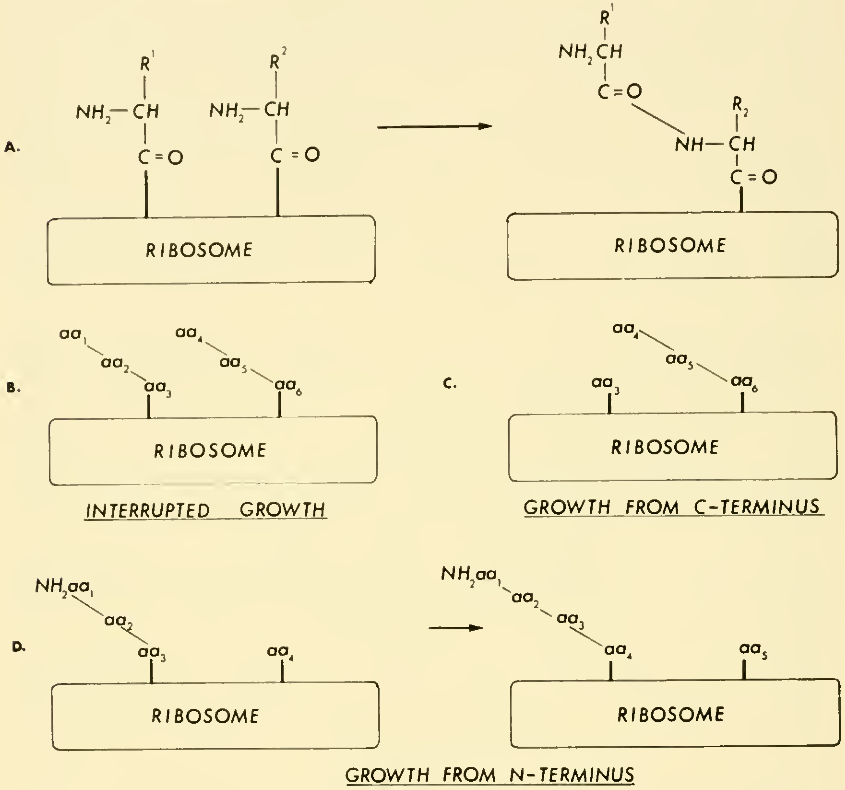

A.
R¹
NH₂—CH
C=O
R²
NH₂—CH
C=O
RIBOSOME
R¹
NH₂CH
C=O
R₂
NH—CH
C=O
RIBOSOME
B.
aa₁
aa₂
aa₃
aa₄
aa₅
aa₆
RIBOSOME
INTERRUPTED GROWTH
C.
aa₄
aa₅
aa₃
aa₆
RIBOSOME
GROWTH FROM C-TERMINUS
D.
NH₂aa₁
aa₂
aa₃
aa₄
RIBOSOME
NH₂aa₁
aa₂
aa₃
aa₄
aa₅
RIBOSOME
GROWTH FROM N-TERMINUS

FIG. 1. Proposed mechanisms of peptide chain growth in protein synthesis.

2. When the peptide bond is formed, one of the amino acids is then no longer attached to the ribosome (Fig. 1A, on the right). From the previous assumption, if a third amino acid is to be added to the chain, it will attach to the ribosome by its carboxyl group, but then will not be able to link if it attaches to the left of amino acid $R^1$, since $R^1$ is not attached to the ribosome. It can link only by attaching to the right of amino acid $R^2$ and attaching to this amino acid.

Thus, if peptides would start to form at *various* points, at some stage (Fig. 1B), $aa_3$ (attached to the ribosome) would need to form a peptide bond with $aa_4$ of another peptide; according to our assumption this is not possible for steric reasons. As shown in Fig. 1C, if growth of the chain started at the C-terminus, a similar situation would arise. However, if the N-terminal amino acid is attached first by its carboxyl group and initiates chain growth (Fig. 1D), the growing tip of the chain ($aa_3$) is attached and can form a peptide bond with $aa_4$, since both are attached to the ribosome. The new growing tip at $aa_4$ is still attached and when $aa_5$ attaches, further chain growth can continue.

How is an ordered synthesis from the N-terminus arranged for? Presumably, this is a function of the "code" and is accomplished by the mechanism of interaction between messenger RNA (see Section IV) and amino acid-specific transfer RNA. The first requirement for the code in this viewpoint is that it can be transcribed only from a single starting point, corresponding to the N-terminal amino acid. Genetic evidence for such a code has been reported by Crick *et al.* (1961). The mechanism by which the N-terminal amino acid is positioned on the messenger RNA-ribosome template *first* is not known, but several possibilities have been discussed by Goldstein (1962). What is the mechanism by which the code ensures that peptides are not formed anywhere along the chain, but only sequentially from the N-terminus? A simple mechanism for providing this arises from a consideration of the mechanism of interaction between the messenger RNA in the ribosome and the incoming amino acyl-RNA. To find its proper place in the amino acid sequence, a triplet of bases in the T-RNA is supposed to hydrogen bond to a triplet of bases in the messenger RNA (see preceding discussion). This alone, however, is postulated to be too weak an interaction to permit the amino acid to remain on the template long enough to participate in further reactions. If it is assumed that in addition to bonding to the code sequence of the messenger RNA, the T-RNA of the incoming amino acid must bond to the preceding molecule of T-RNA, then sequential growth is provided for. According to this idea each T-RNA molecule, in addition to its unique coding triplet of bases, must contain two common sets of sequences which are complementary. For example, each T-RNA would have possibly an AG-rich region and a UC-rich region at different parts of the chain. AC-rich and GU-rich would work also. The AG-rich and UC-rich regions might be 10 or more nucleotides in length and might or might not be exactly complementary in sequence for the required interactions to take place. According to this hypothesis, the incoming amino acyl-RNA would have to bond both by its triplet code to the messenger RNA and by its UC-rich region to the AG-rich region of the T-RNA on the preceding amino acid for effective amino acid transfer. This would be the one at the growing tip of the chain according to the previous discussion.

Other mechanisms can be proposed by which the incoming amino acid could recognize that the preceding amino acid is present. Whatever mechanism is postulated, however, it appears likely that the important characteristic of chain growth from the N-terminal end is that the incoming amino acid and the preceding one are both on the template. This provides steric conditions suitable for peptide bond formation, and could also ensure uninterrupted chain growth, as discussed above.

### 3. *Final Stages in Protein Synthesis*

The previous discussion has dealt with indirect evidence for a particular sequential type of peptide chain growth. It should be noted that details of the mechanism of peptide bond formation and the intermediates directly involved in this process are as yet unknown. The role for GTP suggested earlier, which is hypothetical as yet, is the first approach to this problem. Skipping over these basic questions, therefore, the next problem concerns the mechanism by which finished proteins are released from the ribosome. The best evidence that some mechanism is needed to release finished proteins from their site of synthesis comes from experiments where "nascent," enzymatically active proteins on ribosomes cannot be removed by simple washing techniques (Kihara *et al.*, 1961; Cowie *et al.*, 1961; Siekevitz and Palade, 1960).

In cell-free systems, Simkin (1958) showed that labeled liver microsomes formed soluble protein when incubated with cell sap and an energy-generating system. In the reticulocyte ribosome system, similar studies demonstrated the formation of labeled, soluble hemoglobin when labeled ribosomes were incubated in the complete system ($C^{12}$-amino acids, energy, and enzymes). It was also shown that the ribosomes were not degraded under these conditions (Morris and Schweet, 1961). However, these experiments are essentially "chase" experiments and do not distinguish between *release* of completed proteins and peptide bond synthesis which results in *synthesis* of completed chains. Soluble enzyme fractions which increase the formation of soluble protein without increasing total incorporation greatly have been termed "releasing" enzymes (Lamfrom, 1961), and it has been suggested that these were species specific. However, if ribosomal-bound proteins were nearly completed, the addition of only a few amino acids might result in the formation of soluble protein. Further discussion and evidence for this viewpoint is given in the next section. Until systems are developed where incorporation into ribosomes is *not* occurring while "release" is being measured, the problem of whether or not specific "releasing" enzymes exist remains open.

When the transfer system (from amino acyl-RNA to protein) is used, since ATP is not required, a better assessment of the energy requirements for the formation of soluble protein may be made. Bishop *et al.* (1961) found that little soluble protein was formed in the transfer system, with GTP as the only nucleoside triphosphate. However, this was probably due to the small amount of $C^{14}$-amino acyl-RNA used in these experiments. Formation of soluble protein using a transfer system was shown by von Ehrenstein and Lipmann (1961). However, since GTP plus a generating system was used, it is likely that ATP was formed. In any event, since chains were being formed, the "release" mechanism per se was not studied in either of these cases. Even the formation of soluble protein from labeled ribosomes with GTP and enzyme in the *absence* of amino acyl-RNA (Table III) does not represent release alone, since it is possible that chain growth from intraribosomal intermediates occurs here. Thus, other than the evidence which suggests that some mechanism is involved in liberating completed proteins from the ribosome, practically nothing is known.

Studies of the mechanism of inhibition of protein synthesis by the antibiotic puromycin have been of interest in this connection. Inhibition of protein synthesis puromycin was first reported by Yarmolinsky and de la Haba (1959). Inhibition of transfer from amino acyl-RNA to ribosomes was considered to be the site of action of puromycin. This seemed likely since puromycin has a structure analogous to the amino acyl-adenosine end of transfer RNA. A direct effect of puromycin on the ribosome resulting in the release of soluble protein has been reported for reticulocyte ribosomes (Morris and Schweet, 1961) and liver ribosomes (Hultin, 1961). It was later shown that release of soluble protein by puromycin was maximal in the absence of added soluble enzyme and energy (Morris *et al.*, 1962). Release of soluble protein occurred at 4°C, although more slowly than at 37°C. Ribosome breakdown could not be detected under these conditions. The protein material released by puromycin in the absence of added enzyme was shown to be incomplete globin chains synthesized from the N-terminal end, confirming earlier conclusions (Section III,B,2).

In the *presence of* soluble enzymes and puromycin, in addition, free amino acids and small peptides are found in the TCA-soluble fraction. These have not been well characterized, but it has been suggested that they may represent earlier, intraribosomal intermediates in globin synthesis (Morris *et al.*, 1962). The inhibition of protein synthesis by puromycin is considered to result from the displacement of incomplete globin chains and the postulated earlier intermediates from the ribosome. In the presence of puromycin, further formation of TCA-precipitable peptides would be blocked by a continuation of the same process. This mechanism would permit a certain amount of peptide bond formation in the presence of puromycin, which might result in the formation of small peptides or earlier intermediates; but once the larger peptide chains in the ribosome were released, no incorporation into TCA-precipitable material would occur. The locus of puromycin action is considered to be at the growing point of the chain, which is interrupted at a specific site, possibly when an aromatic amino acid appears in the sequence. This last result is suggested from the fact that neither the leucine analog of puromycin (Rabinowitz and Fisher, 1962), or the free aminonucleoside (A. Morris and R. Schweet, unpublished data) was active. The former authors have studied the effects of puromycin on intact ascites cells. Incorporation into ribosomes was inhibited, but the formation of labeled soluble protein continued. These results were attributed to release of incomplete proteins. However, no release of protein from the ribosomes was observed with intact reticulocytes in this case (Rabinowitz and Fisher, 1962).

Inhibition of incorporation into protein by puromycin in the transfer system from *E. coli* has been studied (Nathans and Lipmann, 1961; Nathans *et al.*, 1962). These authors observed degradation of amino acyl-RNA in the presence of puromycin. This de-acylation was dependent on ribosomes, GTP, and soluble enzymes. It is possible that the de-acylation may be a secondary phenomenon resulting from the incorporation of the $C^{14}$-amino acid of the amino acyl-RNA into small peptides and other intermediates in the ribosome which are then released in the presence of puromycin. Material which was not free amino acid was indeed noted by Nathans *et al.* (1962). Thus, an effect at the first stages of transfer may reflect the primary action of puromycin on an intraribosomal event, as noted earlier in connection with GTP action, where a cofactor required for *transfer* is believed to act on an intraribosomal stage.

To conclude this section on "release," it should be noted that the discussion has been concerned only with the formation of the peptide chain sequence. Specific disulfide bond and tertiary structure formation is currently considered to follow spontaneously (Crick, 1958). The evidence for this viewpoint has been discussed by Berg (1961).

## IV. Information Transfer in Protein Synthesis

In this section, an attempt will be made to review in some detail the question of how the information in DNA specifies a particular amino acid sequence. That DNA is the controlling factor, with some possible exceptions to be discussed, has become clear (see Berg, 1961, for a recent discussion). The mechanism by which DNA transfers this information is now considered to be via formation of a "messenger" RNA. Since the problem of DNA and messenger RNA *synthesis* is reviewed elsewhere (see Chapters VI and VII), the emphasis here will be on the *function* of DNA and RNA in protein synthesis, particularly in cell-free systems. The question of whether a linear sequence of the bases in DNA corresponds to a linear sequence of amino acids is omitted also. The role of RNA in protein synthesis in cell-free systems will be emphasized, since studies in intact bacteria are reviewed by Volkin (Chapter VI). Finally, results with bacteria are treated separately from those using higher organisms because of the very different picture of DNA function in these two cases.

### A. INFORMATION TRANSFER IN BACTERIAL PROTEIN SYNTHESIS

The evidence that DNA carries the information for protein synthesis in most systems (RNA viruses may be an exception) is clear. The question to be considered here is the mechanism involved. One possibility might be that DNA is itself the template for protein synthesis. This does not appear to be the case, although in intact bacteria the evidence is not particularly convincing. Rather, the ability of enucleated algae and amebae to continue protein synthesis has provided the major evidence for this view (see next section). The best evidence that DNA is not the template for protein synthesis in bacteria is the finding that DNA acts to make a particular type of RNA, which *does* appear to participate in the formation of the template for protein synthesis, as discussed below. These studies involve cell-free systems and it must be noted that the evidence using intact bacteria does not distinguish clearly between the direct action of DNA as a template, or the accepted hypothesis that DNA produces an unstable intermediate, the messenger RNA.

The evidence which has led to the present concept of an unstable "messenger" RNA as the intermediate between DNA and protein comes largely from studies of induced enzyme formation (Jacob and Monod, 1961) and phage infection (see Chapter VI). The existence of RNA which is rapidly labeled in intact cells, has the composition of DNA and can form what seem to be specific hybrids with DNA has been demonstrated. These studies have relied largely on labeling experiments and the isolation and purification of messenger RNA has not been achieved, although the report of Bautz and Hall (1962) suggests that this may be accomplished in the near future. However, crude messenger RNA isolated from the intact cell is heterogeneous with regard to molecular weight (see Chapter VI for detailed discussion). This heterogeneity may reflect a structural heterogeneity related to the particular protein for which the RNA contains the sequence information. The separation of the postulated single protein-specific messenger RNA has not been achieved. Further, RNA with the properties described above may be heterogeneous with regard to function and need not consist solely of RNA carrying sequence information for protein synthesis (see Hayashi and Spiegelman, 1961, for a discussion of this point). The synthesis of RNA with the base composition of a primer DNA has been demonstrated in cell-free systems (Weiss, 1960; Weiss and Nakamoto, 1961; Hurwitz *et al.*, 1960; Stevens, 1960). The use of these "RNA polymerase" enzymes in cell-free systems has provided the most powerful evidence for the messenger RNA hypothesis.

### 1. *Functioning of Messenger RNA in Intact Cells*

Among the first clues which suggested that RNA could carry information for protein synthesis was the finding of infective TMV-RNA (Gierer and Schramm, 1956; Fraenkel-Conrat, 1956). At the time, the idea that this infective RNA played a direct role in the synthesis of TMV coat protein and other proteins was not seriously considered. Recent results in cell-free systems (see below) and the effect of mutagenic agents on TMV-RNA (see Chapter X) suggest that TMV-RNA is or contains messenger RNA. However, the molecular weight of infective TMV is higher than the reported values for the sedimentation constants of messenger RNA, which appear to range from 8 to 16S.[5] The size of the minimum unit of TMV which will act as messenger in cell-free systems is not known.

The other systems which have led to the messenger RNA concept also involve "new" protein synthesis. The very careful studies of Volkin and Astrachan (see Chapter VI) showed that following phage infection of *E. coli* a small fraction of the RNA was labeled when $P^{32}$ was added. This RNA had the properties of messenger RNA, particularly with respect to the base composition which resembled the DNA of the infecting phage. Recently, Volkin has shown that the synthesis of this RNA is essential for phage protein synthesis and, in fact, that some stoichiometry exists between the amount of RNA and protein synthesis. This kind of evidence and similar studies which, for example, show a requirement for uracil for phage-induced enzyme synthesis (Pizer and Cohen, 1961) provide strong evidence that a particular type of RNA must be synthesized prior to phage protein synthesis. However, such studies do not define the exact function of this RNA.

The function of this type of RNA, as a component of the template actually making protein, was indicated by finding such RNA bound to ribosomes. RNA having the characteristics of messenger RNA has been reported in ribosomes of phage infected cells by Nomura *et al.* (1960), Gros *et al.* (1961a); and Brenner *et al.* (1961). A similar ribosome-associated RNA was demonstrated in normal (not phage-infected) *E. coli* by pulse-labeling (Gros *et al.,* 1961a). Evidence for messenger RNA in *Pseudomonas aeruginosa, E. coli,* and *Bacillus megaterium* was reported by Hayashi and Spiegelman (1961). There authors used "step-down" cultures, e.g., transferred from a rich to a limited medium, to inhibit ribosomal RNA synthesis and thus permit a more clear-cut labeling of the rapidly synthesized RNA fraction. This fraction was characterized by hybridization to specific DNA (Hall and Spiegelman, 1961). Finally, the elegant experiments of Brenner *et al.* (1961) have shown that no new ribosomes were synthesized after phage infection, and that the newly synthesized messenger RNA, as well as newly synthesized protein, were associated with old ribosomes (present before phage infection).

---

[5] More recent results indicate the sedimentation constants of messenger RNA to range from 8 to 30S (F. Gros, private communication).

The distribution of messenger RNA, in either normal or infected cells, depends on the magnesium ion concentration. Binding to ribosomes occurs in 0.01 $M$ magnesium ion and most of the highly labeled RNA can be dissociated in $10^{-4}$ $M$ magnesium ion. In addition to this type of binding noted by several authors, Gros *et al.* (1961a) found a part of the messenger RNA bound to 70S ribosomes in a more stable linkage. Thus, two stages in messenger RNA function may be, first, a dissociable binding to ribosomes, followed by a more stable linkage. The results of Tissières *et al.* (1960) suggest that ribosomes containing the messenger bound in stable linkage are active in protein synthesis. The results of Brenner *et al.* (1961) are in agreement with this conclusion.

Perhaps the earliest complete "expression" of a messenger RNA concept was that of Jacob and Monod (1961). The concept was based partly on studies mentioned above, but also on studies of induced enzyme formation. While this latter subject is outside the scope of this review, it is not possible to completely avoid this topic in any discussion of messenger RNA. The experiments of Riley *et al.* (1960) illustrate the type of evidence which led to the messenger RNA concept. The structural gene ($z^+$) for $\beta$-galactosidase was introduced into a recipient cell by mating, and within a few minutes the maximal rate of enzyme synthesis of the zygote was attained and continued at a uniform rate. If DNA itself is not the template, then a rapid transfer of information to the enzyme-forming site has taken place. This is similar to the entry of phage DNA into the bacterial cell noted above. However, in addition, the kinetics indicated that a constant number of active templates were present, otherwise the rate of enzyme synthesis would increase with time. Two explanations are considered: (*1*) stable templates are formed, but the gene becomes inoperative after a few minutes of template formation; (*2*) unstable templates are formed which break down and are reformed at some constant rate. Using $P^{32}$-labeled donor bacteria carrying the $z^+$ gene, it was found that breakdown of the genome due to $P^{32}$ decay, as originally noted by McFall *et al.* (1958), led to decreases in β-galactosidase synthesis. These experiments appear to support the second possibility. The concept of a rapidly formed intermediate which is the carrier of information from DNA in the case of induced enzyme formation is supported by many types of evidence. However, that this information is messenger RNA, or RNA at all, has not been shown, in contrast to the case of phage infection. Gros *et al.* (1961b) have reported small increases in RNA labeling on induction, which may represent such synthesis. The evidence that an intact gene is required for β-galactosidase synthesis based on $P^{32}$ decay cannot be interpreted unequivocally, particularly since McFall (1961) has suggested that this effect results from repression of enzyme synthesis by catabolites which accumulate in slowly growing non-viable cells. Thus, the argument for an unstable template which needs to be re-formed by continuous gene action, via messenger RNA synthesis, rests largely on genetic studies and the kinetics of enzyme formation after introduction of DNA into the cell. Another powerful argument for the messenger RNA hypothesis in induced enzyme synthesis is that the mechanism of induction is explained on the same basis, e.g., the template is unstable and needs RNA synthesis continually; in the absence of inducer, this synthesis is repressed. Thus, the inducer acts at the genetic level to derepress messenger RNA synthesis (see Jacob and Monod, 1961, for details of this hypothesis).

In addition to the evidence given above, it has long been known that RNA synthesis is needed for protein synthesis in bacteria. In part, these studies have involved measurement of protein synthesis using purine- and pyrimidine-requiring mutants. If RNA synthesis was limited by the absence of the required base, protein synthesis was depressed (see Berg, 1961, for discussion). The interpretation of this RNA-protein relationship is not clear. Suggestions have been made that common intermediates for RNA and protein synthesis (amino acid-nucleotides) are present. However, more indirect relationships due to poor cell growth (McFall, 1961) may explain this apparent coupling of RNA and protein synthesis (see Spiegelman *et al.*, 1955, for an interesting summary of this problem). More recently, the use of base analogs has permitted additional insight into the RNA-protein relationship. In particular, 5-fluorouracil (see below), azaguanine (Chantrenne, 1959), and thiouracil have been used. Naono and Gros (1960a) and Bussard *et al.* (1960) have studied the synthesis of β-galactosidase and alkaline phosphatase in *E. coli* grown in the presence of the analog, 5-fluorouracil. Very soon after the introduction of the analog, the cells began to synthesize altered enzymes which could be characterized by their immunological and enzymatic properties (β-galactosidase) or by thermolability (alkaline phosphatase). The change from the synthesis of normal to the synthesis of altered enzyme was abrupt and apparently complete. These results were explained by the hypothesis of Chantrenne (1959) that certain nucleic acid analogs may be incorporated into the RNA of the growing cell and exert their influence on protein synthesis by altering the specificity of the RNA. Naono and Gros (1960a; see also Gros, 1960) have observed that, within 1 minute of the addition of the analog, the rates of incorporation of tyrosine and proline into protein dropped while the rate of incorporation of arginine increased. An explanation of this was found in changes in the saturation levels of these amino acids in the amino acyl-RNA pool. Gros and Naono (1961), on the other hand, have found *additional* changes in another RNA fraction. After incubation of bacteria in the presence of $C^{14}$-5-fluorouracil, the radioactivity was found exclusively in a fraction with the sedimentation coefficient 12–14S. After phenol extraction of the RNA, the radioactivity sedimented at 8–10S. The suddenness and completeness of the effect provide good evidence for a rapid synthesis of the RNA fraction involved, and the sedimentation data suggest identity with messenger RNA. Thus, the changes in enzyme synthesis are explained by the synthesis of an altered messenger RNA containing the analog. The problem is somewhat complicated by the finding of incorporation into other types of RNA at later time periods (see Chantrenne and Devreux, 1960, for example) with azaguanine, and the variation in reversal of inhibition when guanine is added after azaguanine incubation for various times.

The final topic for discussion in this section is the "instability" of the messenger RNA. Two types of instability, which may be related, can be distinguished. The first concerns the question of whether messenger RNA breaks down to the nucleotide level before assimilation into ribosomal RNA, or is used at the polynucleotide level. This question is discussed in detail by Roberts (Chapter VII) and by Volkin (Chapter VI). A few observations will be made here. The RNA fraction under consideration in a particular experiment may vary. When, for example, the *criterion* for messenger RNA, is that of a rapidly labeled RNA fraction, this may contain various types of RNA. Part of this may be RNA destined for ribosome synthesis; another part may be messenger (defined as carrying information for protein sequence determination).

Second, the type of experiment may, in fact, determine the fate of the RNA. For example, in phage-infected cells, where nucleotides are used for synthesis of deoxynucleotides, and ribosome synthesis is inhibited, the breakdown of messenger RNA may be emphasized over other pathways. These points are mentioned because of the apparent disagreement in the literature on this question. In any event, it seems that under some conditions there is extensive breakdown of the labeled RNA fraction. In some cases, the specific RNA fraction is loosely bound to ribosomes, but another part is more firmly incorporated in ribosomes which have been described as "active" 70S (and also probably 100S) ribosomes (see below). This RNA is not removed by washing in $10^{-4}\,M$ magnesium chloride (Gros *et al.*, 1961b) and may be the functionally important form of messenger RNA. Finally, a third form may involve synthetic processes which result in the labeled RNA being changed to ribosomal RNA as regards base composition. In this situation it may be functionally inactive.

The important question with regard to functional stability of messenger RNA is related to the mechanism by which it controls protein synthesis. It has been suggested (Gros *et al.*, 1961b) that there is a stoichiometric breakdown of messenger RNA during protein synthesis. This would result if RNA breakdown were directly coupled to peptide bond synthesis. There is little evidence to support such a mechanism. Particularly, in reticulocytes (Bishop *et al.*, 1961, and see below) the stability of messenger RNA rules out such a mechanism. Even in bacteria, there is evidence that a considerable variation in the functional stability of messenger RNA exists. The extreme instability in the case of phage infection and induced enzyme synthesis has been noted. However, even in phage infection, at a later stage the synthesis of messenger RNA is low, while phage synthesis continues (Astrachan and Volkin, 1958).

### 2. *Function of DNA and RNA in Cell-Free Systems*

The cell-free system for incorporation of amino acids developed by Lamborg and Zamecnik (1960) was the first ribosomal system from bacteria. This cell-free system was similar in many respects to cell-free systems from animal tissues. A remarkable property of this system, not found in animal systems (see below), was inhibition by DNase. Tissières *et al.* (1960) obtained up to 70% inhibition of amino acid incorporation with DNase. RNase contamination was excluded and polyelectrolytes did not reverse the inhibition, suggesting that removal of DNA was the cause of inhibition. One of the synthetic compounds used in an attempt to restore activity was poly-A, which was ineffective. Inhibition of cell-free synthesis of β-galactosidase by DNase was reported by Kameyama and Novelli (1960). Complete inhibition was obtained and, in fact, the activity of the initial β-galactosidase present was decreased. A thorough investigation of the DNase inhibition (Matthaei and Nirenberg, 1961) showed that decreases in protein synthesis was noted only after 10–15 minutes of incorporation. The amounts of RNase and trypsin contamination of the DNase preparation were estimated and these amounts shown not to inhibit protein synthesis. The products of DNase digestion did not account for the inhibition. Thus, a requirement for DNA for protein synthesis was demonstrated. The effect of DNA was probably not direct, since complete inhibition (particularly in the initial phases) was not attained. This was attributed to the presence of endogenous messenger RNA which was utilized in the early phases of amino acid incorporation and then destroyed, leading to a complete dependence on DNA at later stages.

Restoration of activity for protein synthesis by RNA (Nirenberg and Matthaei, 1961) as expected was the next step forward. For these studies, the authors used the crude *E. coli* supernatant containing ribosomes, which had been preincubated with DNase and energy. The use of the energy-generating system resulted in incorporation of $C^{12}$-amino acids which presumably led to loss of endogenous messenger RNA. This latter point has not been demonstrated, but, in any event, the treatment provided a system which showed almost no amino acid incorporation in the absence of added messenger RNA. Increases in amino acid incorporation by addition of *E. coli* and yeast ribosomal RNA and by TMV-RNA were observed. The most striking increase was that given by TMV-RNA, which increased incorporation from 42 to 872 counts per minute/mg of ribosomal protein. The RNA-stimulated incorporation was inhibited by puromycin, required energy, and in general showed the properties of the usual amino acid-incorporating systems. Excess soluble RNA was present and did not account for the observed stimulations. Finally, in this paper, stimulation of phenylalanine incorporation by polyuridylic acid was observed (Nirenberg and Matthaei, 1961). These results represent the most significant advance in the understanding of information transfer in protein synthesis in recent years and many new fields are open to investigation as a result. However, the RNA types which gave increased incorporation did not resemble messenger RNA, as defined earlier, and in fact these results did not prove that DNA destruction was involved in the loss of activity.

Tissières and Hopkins (1961) were able to remove DNA from an *E. coli* system by precipitation with magnesium chloride. By adding back *E. coli* DNA plus spermidine, some stimulation of amino acid incorporation was observed. This stimulation required addition of all four nucleoside triphosphates. RNA with the sedimentation properties of messenger RNA was shown to be synthesized. In addition, these authors isolated an RNA fraction, by washing ribosomes in dilute magnesium chloride, which also stimulated amino acid incorporation. This fraction also had the sedimentation characteristics of messenger RNA. These results then suggest that amino acid incorporation depends on the DNA-directed synthesis of messenger RNA and also that a reserve of the latter is present in the cell-free system.

Wood and Berg (1962) removed nucleic acids from an *E. coli* extract by ammonium sulfate and protamine fractionation. Partially purified soluble enzymes and ribosomes, when incubated with the usual components, including added soluble RNA, showed low amino acid incorporation. DNase inhibited only slightly, and the incorporation appeared similar to the residual incorporation found by others to occur in the presence of DNase, e.g., due to endogenous messenger RNA. T2 phage DNA and purified RNA polymerase added to the system gave a 20-fold increase in incorporation. That the DNA effect was due to synthesis of RNA was shown by a number of experiments. These involved two-stage experiments. In the first stage, RNA synthesis was allowed to proceed, or prevented, for example, by omission of UTP. This reaction was stopped and then protein synthesis measured using the previous reaction mixture as the stimulating component. The results showed that RNA synthesized under the influence of DNA was responsible for increased amino acid incorporation, although more striking increases were found when RNA was continuously generated in the amino acid-incorporating system. Since present evidence indicates that RNA polymerase synthesized messenger RNA, it is likely that it is this type of RNA which is effective. However, as in other studies in this field, final proof of this hypothesis will require the demonstration that known, specific proteins are synthesized under the influence of a specific DNA or RNA. Such experiments appear feasible and are in progress in several laboratories. An interesting aspect of these studies was the finding that phage DNA was 5 times more active in stimulating incorporation than *E. coli* DNA, and that heated phage DNA and single-stranded $\phi$X174 phage DNA were inactive. These latter DNA's are efficient primers for the RNA polymerase, and unless the RNA remained bound and unavailable to the protein-synthesizing system, it is difficult to understand the lack of stimulation. The general concept of the role of DNA as primer in the RNA polymerase reaction is that each strand of DNA produces a complementary strand of RNA, resulting in production of RNA of the same over-all composition as the DNA primer. With single-strand

DNA or synthetic deoxypolymers, the complementary RNA is produced (see Chamberlin and Berg, 1962, for references and recent results). Bautz and Hall (1962) have suggested that *in the intact cell* only one of the two strands of DNA is active, since the T4-specific RNA they isolated did not have the composition of the DNA. The solution to this question is critical to the understanding of the mechanism of information transfer and it is likely that these answers will come from further studies of DNA-directed protein synthesis in cell-free systems.

Studies of the cell-free synthesis of a specific protein, β-galactosidase, were reported by Kameyama and Novelli (1960). This cell-free system from *E. coli* produced a net increase in β-galactosidase activity. Addition of DNase resulted in complete inhibition of enzyme synthesis, accompanied by an 85% inhibition of amino acid incorporation into total protein (Novelli *et al.*, 1962). After X-irradiation of *E. coli* synthesizing β-galactosidase constitutively, the combined ribosomal and supernatant fractions were unable to synthesize β-galactosidase. Addition of DNA isolated from cells constitutive for β-galactosidase synthesis, or from induced cells, repaired the lesion introduced by X-irradiation. DNA isolated from non-inducible cells was ineffective, providing a control for non-specific effects of added DNA. In addition, in order for DNA from inducible cells to stimulate β-galactosidase synthesis, the DNA must be prepared from preinduced cells. This finding would be consistent with the concept of repressors acting at the gene level (Jacob and Monod, 1961). However, the ribosomes used must also come from preinduced cells, which suggests perhaps that specific ribosomes are synthesized during preinduction. The results discussed previously have indicated a lack of specificity of the ribosome and, in fact, suggest that all types of proteins can be made in any population of ribosomes depending on the messenger RNA provided. Even when all the components of the system for β-galactosidase synthesis came from preinduced cells, inducer was required for cell-free synthesis of β-galactosidase (Novelli *et al.*, 1962). This suggested production of repressor in the system under the influence of the $i^+$ gene (Jacob and Monod, 1961). Indeed, when all the constituents came from $i^-$ cells, (constitutive mutant) inducer was not required and DNA from $i^+$ cells, not induced, was inhibitory. This result was interpreted as indicating a binding of repressor to DNA. These results represent a remarkable advance in the understanding of induced enzyme synthesis. They differ enough from other results that confirmation in other laboratories would be desirable. Also, since there is always a considerable amount of preformed enzyme present at the start of the experiment, the usual enzyme synthesis measured is only a doubling of that already present. However, careful immunological and electrophoretic studies by the authors provide good evidence that new enzyme synthesis has occurred in this system.

### 3. *Function of Synthetic Polynucleotides in Cell-Free Systems*

The DNase-treated, preincubated cell-free system from *E. coli,* described earlier, was stimulated by various types of RNA (Nirenberg and Matthaei, 1961). However, it was not clear that this RNA served an informational function, e.g., directed the synthesis of specific proteins by specifying the sequence of amino acids. Using poly-U as the stimulating polynucleotide, however, only phenylalanine was incorporated. Evidence was provided that the product was polyphenylalanine. In this artificial system, therefore, the polynucleotide was specifying which amino acid was incorporated. In a subsequent publication (Nirenberg *et al.,* 1962), it was shown that this system had many of the properties of usual incorporating systems. For example, phenylalanyl-RNA was an intermediate. Other synthetic polynucleotides were then tested in both Nirenberg's and Ochoa's laboratories. The results of these studies have been extremely fruitful, in that only certain polynucleotides stimulated the incorporation of particular amino acids. Using certain assumptions noted below, it was possible to assign particular nucleotide triplets to particular amino acids and thus provide the solution to the code. The findings have been published in some detail; references to these are found in the most recent publications (Basilio *et al.,* 1962; Matthaei *et al.,* 1962).

The primary basis for assignment of code letters rests on the finding that poly-U stimulates only phenylalanine incorporation. In fact, incorporation of leucine and isoleucine were also stimulated, but to a much smaller extent. When polymers containing two or more nucleotides were used, they were assumed to contain random sequences. For example, the U:C polymer containing 5 times as much U as C was considered to contain 5 times as many UUU triplets as UUC, UCU, or CUU triplets. The use of a triplet is indicated as a minimum number for each coding unit by the finding that incorporation of some amino acids was stimulated only by polymers containing three different nucleotides. The results of Crick *et al.* (1961) (see Chapter VI) also indicate a triplet code. The term $U_2C$ includes all the possible arrangements of these three letters; thus the sequence is not known for any code letter. The letters in the code word are then decided by comparing the ratio of phenylalanine stimulation to the amino acid in question. For example, using the case given above, the only amino acids (besides phenylalanine) whose incorporation was stimulated by poly-UC were leucine, proline, and serine (Speyer *et al.,* 1962). The ratios were 5, 13 and 4, respectively. The code words assigned then were $U_2C$ for leucine and serine, and $UC_2$ for proline (calculated ratio is 25). However, proline is also stimulated by a UAC polymer. The phenylalanine/proline ratio here was 29, expected $UUU/UC_2$ in the polymer was 36. Thus, agreement was satisfactory. However, a UCG polymer with the same $UUU/UC_2$ ratio did not stimulate proline incorporation. The results for most of the amino acids do not contain discrepancies of this type. But not many polymers of different types with similar calculated ratios of two triplets have been tested. Another example of this type is tryptophan. The code word assigned by Speyer *et al.* (1962) and Matthaei *et al.* (1962) is $UG_2$. The ratios of stimulated incorporation by UG and UCG polymers agree, but stimulation by UAG was not observed and should have been the same as UCG (Speyer *et al.*, 1962). On the other hand, Matthaei *et al.*, for the same amino acid, but using polymers of different composition, confirm the assignment with UG and UGA, but not with UGC. In any event, a number of amino acids are given the same code letters by both laboratories, although some differences are noted. Certain amino acids, such as glutamic acid, were stimulated only by UAG so this word was assigned although the UUU/UAG ratio did not agree with incorporation ratios.

Despite these problems, which may be explained in the future, some remarkable conclusions can be made from the results, and also new problems are opened for investigation. For example, some of the earlier codes are eliminated. All of the code words contain uridylic acid (see Chapter VI for code words). The question arises whether this is a real phenomenon or results from the technique used. Since it is likely that UUU is the code for phenylalanine, and poly-A and poly-C do not code for any amino acid, it is clear that polymers containing large amounts of U will give runs of phenylalanine, thus providing a matrix for other amino acids to attach to and yield an insoluble peptide. Both laboratories have only looked for insoluble peptides. Therefore, code words containing no U may exist, but would yield stimulation into small peptides, soluble in trichloroacetic acid. Matthaei *et al.* (1962) have noted that poly-AG did not stimulate incorporation, but small peptides were not looked for. If this argument is true then even the failure of poly-A or poly-C to stimulate incorporation may not be correct. There are 37 triplets containing U and all the amino acids are coded for by 24 triplets. There are two words for asparagine, and threonine, and three words for leucine, according to Speyer *et al.* (1962). Matthaei *et al.* (1962) list serine as having two code words. Thus, 12 triplets with U are not accounted for. Either these are "nonsense," or further degeneracies (more than one code word per amino acid) exist. If triplets without U could also code, the degeneracy must be even greater. On the other hand, if all codes contain U, the composition of DNA and messenger RNA does not correspond (does not contain sufficient U) and it must be supposed that part of these polynucleotides does not code, e.g., is nonsense or is used for some other function.

An important impetus in support of the validity of the code letters has come from studies of amino acid replacements in mutant proteins. In particular, the nitrous acid-induced mutants are of interest. They are produced by nitrous acid treatment of infective TMV-RNA. After infection and replication, mutants are isolated and the protein structure determined (see Chapter X). It is considered that most if not all of the changes are due to the action of nitrous acid on adenine and cytosine, which yield guanine and uracil, respectively, in the new RNA. The amino acid changes being known, it is possible to compare these with the change in the code words derived from cell-free systems. The data have been compiled from the literature in both laboratories and the agreement is good. Speyer *et al.* (1962) record that 11 of 16 replacements agree. For example, a common mutation results in threonine $\dashrightarrow$ isoleucine. The codes are UAC $\dashrightarrow$ UAU, which is a C to U change and agrees with the action known for nitrous acid. Important exceptions are the aspartic acid $\dashrightarrow$ serine, and serine $\dashrightarrow$ leucine replacements, which involve either change of two bases or an A $\dashrightarrow$ U change based on $U_2C$ for serine. However, if a second code word for serine were UCG, some of these discrepancies would disappear. Matthaei *et al.* (1962) include UCG as a second code word for serine, but their published data do not lend strong support to this assignment. In general, then, these results support the proposed code words, since the combined effect of the known action of nitrous acid and the proposed code words are so restrictive to the possible amino acid changes that the agreement found must have significance.

This is reinforced by the conclusions reached by Smith (1962) in a study of amino acid replacements in human hemoglobins. The question studied is whether known amino acid changes in the abnormal hemoglobins are compatible with the change of a single base in the synthetic code words. In this case there are no restrictions to the kind of base change involved and in fact almost every type must occur to account for the amino acid changes. Smith predicted the code words for four amino acids based on known ones. For example, glutamic acid had not yet been coded; in fact, the data with synthetic polynucleotides is still weak for this amino acid (Speyer *et al.*, 1962). But since glutamic acid could be replaced by valine ($U_2G$), lysine ($UA_2$), and glycine ($UG_2$) in various hemoglobins, the code word for glutamic acid must contain

U, A, and G. This is consistent with another hemoglobin where glutamic acid is replaced by glutamine, which is UGC. The code word for glutamine, UGC, was predicted from TMV replacement data, but so far the synthetic polynucleotide, UGC, has not stimulated glutamine incorporation. Smith also considered the question of sequences of the bases in any code word, and it seems likely that U may only occupy two of the three possible positions. Further, if the sequence of one code word containing three bases is established, most of the others can be deduced. Amino acid replacements in proteins of various species were also compared, and here the situation was not as clear. Some of the changes must involve two mutations if the code letters are correct. However, this may not be too unusual. The implications of these comparisons in the evolution of protein structure and function are of great interest.

One of the important conclusions from the agreement in code words derived from amino acid replacements and those derived from stimulation of amino acid incorporation is that the code is "universal." That is, in general the code word for an amino acid is the same in mammals, *E. coli*, and TMV. A similar conclusion was reached from the finding that transfer RNA from various species could be used to synthesize hemoglobin using reticulocyte ribosomes. This conclusion does not exclude differences in composition or specificity in messenger RNA from different species, but would ascribe these to non-coding portions of the molecule. Nonetheless, it is still important to show that poly-U will code only for phenylalanine using mammalian ribosomes,[6] for example, or that a particular messenger RNA directs the synthesis of the same protein in ribosomes of two different species.

This question is also related to the question of the role of the ribosome in protein synthesis. While ribosomes of *E. coli* can accept information from different types of natural and synthetic messengers, only 10% or less of the ribosomes contain labeled poly-U when this is used as messenger. The rest of the labeled poly-U is rapidly degraded, but phenylalanine incorporation continues, implicating these special ribosomes as the "active" sites of protein synthesis (Matthaei *et al.*, 1962; M. Nirenberg, private communication). Do these "active" ribosomes represent a special class, or are they representative of all the ribosomes of the cell? Before discussing this question, it is of interest to note that these data provide evidence for a catalytic role of the messenger RNA in the active ribosomes. Since one phenylalanine residue is incorporated per uridylic acid residue under favorable conditions, and much of the poly-U is probably never integrated into functional ribosomes, plus the fact that three U's code for a single amino acid, it is likely that the poly-U messenger acts catalytically, even in the cell-free system.

---

[6] This result has now been reported by H. Arnstein (*Nature*, in press) and confirmed in this laboratory.

The data shown in Table V indicate that poly-U or TMV-RNA stimulated ribosomes synthesize protein in similar amounts to the original crude extracts. The somewhat lower values are undoubtedly due in part to the manipulations involved in removing DNA and endogenous messenger RNA from the artificial systems before assay. If we assume that ribosomes stimulated by poly-U or other added messenger do not incorporate more amino acids per ribosome than in the crude extract, then these data indicate that most of the ribosomes which were making many proteins in the crude extract are now making only polyphenylalanine or a few proteins involved in TMV synthesis. Thus, from the information transfer standpoint there seems to be little ribosome specificity. A point which will be returned to is the fact that reticulocyte ribosomes, which contain a stable messenger RNA, incorporate amounts of amino acids that are similar to the stimulated bacterial cell-free system (Table V).

TABLE V
INCORPORATION OF AMINO ACIDS IN CELL-FREE SYSTEMS

| System | Amino acid | Incorporation[a] | Total[a] protein | Reference |
|---|---|---|---|---|
| TMV-RNA | Valine | 0.4 | 4.5 | Nirenberg and Matthaei (1961) |
| Poly-U | Phenylalanine | 6 | 6 | Nirenberg and Matthaei (1961) |
| T2 phage DNA and RNA polymerase | Leucine plus valine | 0.7 | 5 | Wood and Berg (1962) |
| Poly-U | Phenylalanine | 24.6 | 24.6 | Lengyel *et al.* (1961). |
| Original crude extract | Individual amino acids | — | 50 | Tissières *et al.* (1960) |
| Original crude extract | Valine | 1.0 | 18 | Matthaei and Nirenberg (1961) |
| Reticulocyte ribosomes | Leucine | 1.5 | 11 | Schweet *et al.* (1961) |

[a] Values are given as mμmoles of amino acid per mg of ribosomal protein.

Ribosome specificity has, however, been noted by Novelli *et al.* (1962). Also, as noted earlier, there is an enzymatic specificity, since transfer enzymes from rabbit reticulocytes and *E. coli* were not interchangeable (Nathans *et al.,* 1962). Since it appears that transfer and messenger RNA are not species specific, it follows that the enzymatic specificity serves another function, not involving the template (which is considered to be the ribosome-messenger RNA region with which the transfer RNA interacts). This function could involve intermediates in the ribosome beyond the amino acyl-RNA stage as noted earlier.

### B. INFORMATION TRANSFER IN NUCLEATED ORGANISMS

In this section again, the main emphasis will be on studies with cell-free systems. Since the nucleus is the site of the hereditary control of protein specificity in nucleated organisms, we shall concern ourselves here with evidence and speculation regarding the method by which the nucleus exerts its control. The controlling role of the nucleus, long a subject of speculation, was first clearly indicated by the work of Hämmerling (1953) with the alga *Acetabularia*. These experiments, experiments with enucleated *Amoeba* (reviewed by Brachet and Chantrenne, 1956) and human amnion cells (Goldstein *et al.*, 1960), demonstrate (1) that protein synthesis can continue for extended, *but not indefinite* periods of time in the absence of the nucleus, (2) that some long-term control, particularly of differentiation (as in the maturation of *Acetabularia*), is under nuclear control. It is impossible, from such experiments, to decide whether any fraction of protein synthesis is *not* under long-term nuclear control. Furthermore, it is not clear at the moment whether protein synthesis in enucleated organisms might not be under the control of cytoplasmic DNA.

#### 1. *Synthesis of Ribosomes*

Evidence concerning the site of synthesis of the ribosomes is indeed scanty. Much of the RNA of the nucleus appears to be located in ribonucleoprotein particles. In pea seedling nuclei, for example, 20% of the RNA is found in ribosomes (Rho *et al.*, 1961). Furthermore, the ribosomes of thymus nuclei are active in protein synthesis (Frenster *et al.*, 1960) by pathways which are essentially the same as those followed in cytoplasmic synthesis (reviewed by Allfrey and Mirsky, 1961). It is not clear, however, whether the nuclear ribosomes are in the process of synthesis and are destined to be passed to the cytoplasm, or are engaged in the synthesis of the enzymes and other proteins found within the nucleus.

Using pea seedlings, Ts'o and Sato (1959) have observed that, after short periods of labeling, the RNA of the nuclear ribosomes was 50–100 times more radioactive than the RNA of the cytoplasmic ribosomes. If this is not due to a phenomenally rapid synthesis of ribosomes of nuclear function, it indicates that the ribosomes are synthesized in the nucleus. Schneider (1961) has observed a transfer of radioactivity from the RNA of prelabeled ($C^{14}$-orotic acid) thymus nuclei to unlabeled cytoplasm upon incubation of the two fractions together. The labeled RNA in the cytoplasm was found associated with microsomal and ribosomal fractions. It is clear that intact RNA molecules were transferred from nucleus to cytoplasm, but whether they were incorporated into the ribosomes before or after transfer is not clear. These two shreds of evidence might *suggest* a nuclear origin of cytoplasmic ribosomes.

### 2. *Synthesis of RNA*

Early enucleation studies with *Acetabularia* and *Amoeba* (Brachet, 1955) indicated that RNA synthesis was impaired in the absence of a nucleus. The more recent work of Prescott (1960) and Goldstein *et al.* (1960) indicates that the inhibition of RNA synthesis may be complete in enucleated *Amoeba*, *Acanthamoeba*, and human amnion cells. Recent autoradiographic experiments (Perry *et al.*, 1961a) are indicative of more specific effects. After inactivation of the nucleolus with an ultraviolet microbeam, incorporation of cytidine into cytoplasmic RNA was severely inhibited, while incorporation into the nucleus (apart from the nucleolus) was inhibited to a lesser extent. In this case, too, however (Perry *et al.*, 1961b) protein synthesis (as measured by $H^3$-lysine incorporation) was inhibited, although to a lesser extent than RNA synthesis.

Nuclear transplantation studies in *Amoeba* (Goldstein and Plaut, 1955) showed that a transfer of RNA from nucleus to cytoplasm occurred. The work of Schneider (1961) with isolated nuclei is in agreement with this. Kinetic studies of RNA synthesis in HeLa cells (Feinendegen *et al.*, 1961) and hyphae of *Neurospora crassa* (Zalokar, 1960) indicate that by far the greater part of RNA synthesis in intact cells occurs in the nuclei. The elegant studies of Zalokar (1960) using *Neurospora* hyphae centrifuged to separate the different elements of cellular organization, *without cell rupture*, indicate that the RNA transferred from the nucleus to the cytoplasm becomes associated with the reticulum, presumably with the ribosomes. Thus a nuclear origin of ribosomal RNA is suggested, but by no means proven. The cytoplasmic synthesis of a metabolically stable ribosomal RNA would presumably pass undetected in these studies.

Rho and Bonner (1961) have fractionated pea seedling nuclei prelabeled with RNA precursors, and have concluded that the primary site of RNA synthesis is the chromatin. The newly synthesized RNA was transferred to the nucleolus. Some autoradiographic studies are in agreement, while others are not. Reference to this work will be found in the paper of Rho and Bonner (1961).

The type of RNA synthesized in these studies is not known. In many of the autoradiographic studies, it is likely that the synthesis involved ribosomal RNA, based on the large amount of synthesis. The finding of a rapidly labeled RNA with the base composition of DNA (Sibatani *et al.,* 1962) in thymus nuclei is evidence that "informational" RNA can be synthesized in the nucleus. The demonstration of an RNA polymerase enzyme in thymus nuclei (Weiss, 1960) is consistent with this viewpoint. The association of newly synthesized RNA with chromatin DNA (Bonner *et al.,* 1961) using pea embryo preparations is also suggestive of this type of RNA, which may also be a precursor of ribosomal RNA. However, it is not yet known whether messenger RNA synthesis occurs *exclusively* in the nucleus, or also to a small extent in the cytoplasm. In addition, is there destruction of messenger RNA and replacement into old ribosomes from nuclear messenger RNA as in bacteria? Or is the complete package made in the nucleus? The continued synthesis of protein in the absence of a nucleus and of net RNA synthesis (see Goldstein *et al.,* 1960, for references), does not exclude the presence of a small fraction of RNA with a rapid turnover, presumably synthesized under the direction of cytoplasmic DNA. The alternative is a stable messenger RNA integrated into the ribosome. During the lifetime of a nucleated cell, depending on the stage of differentiation, it is possible that both mechanisms play a role.

The messenger RNA hypothesis (Jacob and Monod, 1961) is based on phenomena observed in bacteria, and we have seen that, though no rigorous proof is available, the circumstantial evidence is largely consistent with it. Rapid adaptation is a characteristic, and indeed necessary, property of bacteria. On the other hand, although adaptation is not unknown in cells of higher organisms, the changes which normally occur are slow and often irreversible. Thus, information turnover might differ in rate and even in basic mechanism from that in bacteria.

### 3. *Studies with Cell-Free Systems*

Studies of information transfer with cell-free systems of higher organisms have been done mainly with the reticulocyte system. This type of cell represents an extreme degree of differentiation. Measurements of $P^{32}$ and guanosine incorporation into RNA of *intact* reticulocytes showed a turnover of RNA too low to be compatible with messenger RNA formation during hemoglobin synthesis (Nathans *et al.,* 1962). Similar results were obtained in this laboratory. For example, after 2 hours of incubation of intact reticulocytes, the incorporation of either $P^{32}$- or $H^3$-cytidine into total RNA was between 0.1 to 0.5 m$\mu$moles per ml of packed cells, while amino acid incorporation was 6,000 to 10,000 m$\mu$moles. (G. Favelukes and R. Schweet, unpublished data). Thus, in this extreme case, assuming the experiments are testing RNA turnover adequately, there is no evidence for any kind of information turnover.

The studies with cell-free systems have attempted to study this question further, and also to change the "information" in reticulocyte ribosomes, even though this might not occur in the intact cell. Lamfrom (1961) incubated the ribosomes of rabbit or sheep reticulocytes with enzyme fractions from rabbit or sheep supernatants, and concluded that the hemoglobin type synthesized was partly a function of the species from which the ribosomes were derived. However, some hemoglobin corresponding to the species from which the soluble fraction was derived was also found. It was concluded that the pH 5 precipitate was inactive in this respect, and that the ability to change the specificity of the ribosomes was found in the pH 5 supernatant. Kruh *et al.* (1961) described a rather different situation. Using ribosomes and soluble fractions derived from the reticulocytes of the rabbit and the guinea pig, considerable synthesis of the heterologous type of hemoglobin was observed when the ribosomes of one species were incubated with the pH 5 precipitate of the other. Thus, Kruh *et al.* (1961) concluded that soluble fractions could change the type of hemoglobin synthesized, but the opposite enzyme fraction was effective. Bishop *et al.* (1961), using reticulocytes of the rabbit and the mouse, incubated the ribosomes of one species with a mixture of the ribosomal supernatants of *both* species, and found that the hemoglobin synthesized corresponded to the species which provided the ribosomes. It is possible that the use of mixed supernatants resulted in the failure of a potential specificity change to be expressed if greater affinity existed between homologous than between heterologous factors. Recently, Miller and Lamfrom (1962) extended these studies to duck-rabbit mixtures. In this case, hemoglobin corresponding to the *ribosome* species was synthesized, and in addition radioactivity was found in a protein peak which appeared to be a hybrid hemoglobin (composed of chains from both species). Thus, the evidence for the synthesis of hemoglobin different from the ribosome species used is conflicting, and further studies are required to decide whether the addition of supernatant fractions from one species can result in synthesis of the hemoglobin of that species when ribosomes of another species are used. In addition, no evidence that RNA is involved has been presented as yet.

Another approach to the problem of information transfer has been to study whether it is *necessary* to provide sequence information to the ribosome in order to synthesize hemoglobin in the cell-free system. Both Kruh *et al.* (1961) and Lamfrom (1961) reported that no soluble hemoglobin was formed when reticulocyte ribosomes were incubated with pH 5 enzymes from livers of various animals. Kruh *et al.* (1961) suggested that this is a failure to form the peptide chain, while Lamfrom (1961) postulated that species-specific "releasing" enzymes might be involved. Bishop and Schweet (1962) studied this question using washed reticulocyte ribosomes and guinea pig liver enzymes. The results with pH 5 enzyme fractions were similar to those noted above; little labeled, soluble hemoglobin was formed. However, this failure did not involve species specificity, but was related to the *low incorporation* observed with this enzyme fraction. When a soluble fraction from liver purified by ammonium sulfate fractionation was used, soluble, labeled hemoglobin was formed. The soluble fraction used was prepared from a liver supernatant as previously described for reticulocyte enzymes (Allen and Schweet, 1962). The amount of labeled hemoglobin formed with liver enzymes plus reticulocyte ribosomes was similar to the amount formed with reticulocyte enzymes added at a level to give the same total incorporation.

TABLE VI
PROPERTIES OF *Escherichia coli* AND RETICULOCYTE RIBOSOMES[a]

| Treatment | *E. coli* system | Reticulocyte system |
|---|---|---|
| Ribonuclease (1 $\mu$g/ml) | Complete inhibition | Complete inhibition |
| DNase (5–50 $\mu$g/ml) | 55% inhibition | No effect |
| Chloramphenicol ($10^{-3}$ $M$) | 91% inhibition | 3% inhibition |
| Puromycin ($10^{-4}$ $M$) | 95% inhibition | 95% inhibition |
| Add TMV-RNA after DNase | 80-fold stimulation | No effect |
| Internal ribosomal ribonuclease | Present | Absent |

[a] Data summarized from Tissières *et al.* (1960); Nathans *et al.* (1962); Schweet *et al.* (1961); and unpublished data.

Lack of inhibition by DNase (Table VI) is consistent with the concept that information transfer is not necessary for hemoglobin synthesis in the cell-free system as suggested by the lack of RNA synthesis in the intact cell. However, even in the absence of addition of messenger RNA, the activity of the reticulocyte ribosomes is comparable to "stimulated" *E. coli* ribosomes (Table V). Further efforts along these lines in several laboratories involve attempts to add various types of synthetic polynucleotides or messenger RNA, such as TMV (Table VI). So far our own results with TMV-RNA have been negative. However, a striking stimulation of phenylalanine incorporation has been obtained with poly-U, confirming the report by Arnstein.[6] Further, hemoglobin synthesis can be almost completely abolished by pre-incubation in the complete system or by incubation with chelating agents, without affecting the poly-U stimulated incorporation. These results (R. Arlinghaus and R. Schweet, unpublished data) indicate that hemoglobin and polyphenylalanine are synthesized on separate sites.

It seems likely that a fraction similar in function to messenger RNA is present in reticulocyte ribosomes (G. Favelukes, B. Hardesty, and R. Schweet, unpublished data). Ribosomal RNA from reticulocytes stimulated amino acid incorporation when added to the usual preincubated *E. coli* system. Whether this stimulation represents even a small amount of globin synthesis or not, is not known. Our present concept, therefore, is that the basic mechanism of information transfer in higher organisms is similar to bacteria. However, at least in the reticulocyte, the messenger RNA is firmly bound in the ribosome and if it is destroyed and replaced at all, this is a slow process. The synthesis of messenger RNA probably occurs only in the nucleus.

## REFERENCES

Allen, E. H., and Schweet, R. S. (1960). *Biochim. et Biophys. Acta* 39, 185.

Allen, E., and Schweet, R. (1962). *J. Biol. Chem.* 237, 760.

Allen, E., Glassman, E., Cordes, E., and Schweet, R. (1960). *J. Biol. Chem.* 235, 1068.

Allfrey, V. C., and Mirsky, A. E. (1961). *In* "Protein Biosynthesis" (R. J. C. Harris, ed.), pp. 67–81. Academic Press, New York.

Anfinsen, C. B., and Steinberg, D. (1951). *J. Biol. Chem.* 189, 739.

Askonas, B., Campbell, P., Godin, C., and Work, T. (1955). *Biochem. J.* 61, 105.

Astrachan, L., and Volkin, E. (1958). *Biochim. et Biophys. Acta* 29, 536.

Basilio, C., Wahba, A., Lengyel, P., Speyer, J., and Ochoa, S. (1962). *Proc. Natl. Acad. Sci. U.S.* 48, 613.

Bautz, E., and Hall, B. (1962). *Proc. Natl. Acad. Sci. U.S.* 48, 400.

Benzer, S., and Weisblum, B. (1961). *Proc. Natl. Acad. Sci. U.S.* 47, 1149.

Berg, P. (1956). *J. Biol. Chem.* 222, 1025.

Berg, P. (1961). *Ann. Rev. Biochem.* 30, 293.

Berg, P., and Lagerkvist, V. (1962). *J. Biol. Chem.*, in press.

Berg, P., and Ofengand, E. (1958). *Proc. Natl. Acad. Sci. U.S.* 44, 78.

Berg, P., Bergmann, F., Ofengand, E., and Dieckmann, M. (1961). *J. Biol. Chem.* 236, 1726.

Bergmann, F., Berg, P., and Dieckmann, M. (1961). *J. Biol. Chem.* 236, 1735.

Bishop, J., and Schweet, R. (1961a). *Federation Proc.* 20, 389.

Bishop, J., and Schweet, R. (1961b). *Biochim. et Biophys. Acta* 49, 235.

Bishop, J., and Schweet, R. (1961c). *Biochim. et Biophys. Acta* 54, 617.

Bishop, J., and Schweet, R. (1962). *Biochim. et Biophys. Acta,* in press.

Bishop, J., Allen, E., Leahy, J., Morris, A., and Schweet, R. (1960a). *Federation Proc.* 19, 347.

Bishop, J., Leahy, J., and Schweet, R. (1960b). *Proc. Natl. Acad. Sci. U.S.* 46, 1030.

Bishop, J., Favelukes, G., Schweet, R., and Russell, E. (1961). *Nature* 191, 1365.

Bloemendal, H., Bosch, L., and Sluyser, M. (1960). *Biochim. et Biophys. Acta* 41, 454.

Bloemendal, H., Littauer, U. Z., and Daniel, V. (1961). *Biochim. et Biophys. Acta* **51**, 66.

Bonner, J., Huang, R., and Maheshwari, N. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 1548.

Borsook, H. (1952). *Fortschr. Chem. org. Naturstoffe* **9**, 202.

Borsook, H., Deasy, C., Haagen-Smit, A., Keighley, G., and Lowy, P. (1950). *J. Biol. Chem.* **187**, 839.

Bosch, L., Bloemendal, H., and Sluyser, M. (1959). *Biochim. et Biophys. Acta* **34**, 272.

Bosch, L., Bloemendal, H., and Sluyser, M. (1960). *Biochim. et Biophys. Acta* **41**, 444.

Brachet, J. (1955). *Biochim. et Biophys. Acta* **18**, 247.

Brachet, J., and Chantrenne, H. (1956). *Cold Spring Harbor Symposia Quant. Biol.* **21**, 329.

Brenner, S., Jacob, F., and Meselson, M. (1961). *Nature* **190**, 576.

Brown, G., Brown, A., and Gordon, J. (1959). *Brookhaven Symposia in Biol.* **12**, 47.

Bussard, A., Naono, S., Gros, F., and Monod, J. (1960). *Compt. rend. acad. sci.* **250**, 4049.

Campbell, P. N. (1962). *Fifth Intern. Congr. Biochem., Moscow, 1962,* Vol. **2**. In press.

Campbell, P. N. (1961). *Biol. Revs. Cambridge Phil. Soc.* **35**, 413.

Campbell, P. N., Greengard, O., and Kernot, B. A. (1960). *Biochem. J.* **74**, 107.

Cantoni, G. L. (1962). *Strasbourg Conf.* Private communication.

Chamberlin, M., and Berg, P. (1962). *Proc. Natl. Acad. Sci. U.S.* **48**, 81.

Chantrenne, H. (1959). *Biochem. Pharmacol.* **1**, 233.

Chantrenne, H., and Devreux, S. (1960). *Biochim. et Biophys. Acta* **41**, 239.

Cowie, D., Spiegelman, S., Roberts, R., and Duerksen, J. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 114.

Crick, F. H. C. (1958). *Symposia Soc. Exptl. Biol.* **12**, 138.

Crick, F. H. C., Barnett, L., Brenner, S., and Watts-Tobin, R. (1961). *Nature* **192**, 1227.

Dalgliesh, C. E. (1958). *Rec. trav. chim.* **77**, 634.

Dintzis, H. M. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 247.

Dintzis, H. M., Borsook, H., and Vinograd, J. (1958). *In* "Microsomal Particles and Protein Synthesis" (R. B. Roberts, ed.), pp. 95–99. Pergamon Press, New York.

Doctor, B., Apgar, J., and Holley, R. (1961). *J. Biol. Chem.* **236**, 1117.

Dunn, D. B., Smith, J. D., and Spahr, P. F. (1960). *J. Mol. Biol.* **2**, 113.

Feinendegen, L. E., Bond, V. P., and Painter, R. B. (1961). *Exptl. Cell Research* **22**, 381.

Fessenden, J., and Moldave, K. (1961). *Biochem. Biophys. Research Communs.* **6**, 232.

Fraenkel-Conrat, H. (1956). *J. Am. Chem. Soc.* **78**, 882.

Frenster, A., Allfrey, V. C., and Mirsky, A. E. (1960). *Proc. Natl. Acad. Sci. U.S.* **46**, 432.

Gierer, A., and Schramm, G. (1956). *Nature* **177**, 702.

Glassman, E., Allen, E., and Schweet, R. (1958). *J. Am. Chem. Soc.* **80**, 4427.

Goldstein, A. (1962). *J. Mol. Biol.* **4**, 121.

Goldstein, A., and Brown, B. J. (1961). *Biochim. et Biophys. Acta* **53**, 438.

Goldstein, L., and Plaut, W. (1955). *Proc. Natl. Acad. Sci. U.S.* **41**, 874.

Goldstein, L., Micou, J., and Crocker, T. T. (1960). *Biochim. et Biophys. Acta* **45**, 82.

Gros, F. (1960). *In* "The Nucleic Acids" (E. Chargaff and J. N. Davidson, eds.), Vol. III, pp. 409–452. Academic Press, New York.

Gros, F., and Naono, S. (1961). *In* "Protein Biosynthesis" (R. J. C. Harris, ed.), pp. 195–206. Academic Press, New York.

Gros, F., Hiatt, H., Gilbert, W., Kurland, C. G., Risebrough, R. W., and Watson, J. D. (1961a). *Nature* **190**, 581.

Gros, F., Gilbert, W., Hiatt, H., Attardi, G., Spahr, F., and Watson, J. (1961b). *Cold Spring Harbor Symposia Quart. Biol.* **26**, 111.

Grossi, L. G., and Moldave, K. (1960). *J. Biol. Chem.* **235**, 2370.

Hall, B. D., and Spiegelman, S. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 137.

Hämmerling, J. (1953). *Intern. Rev. Cytol.* **2**, 475.

Hartmann, G., and Coy, U. (1961). *Biochim. et Biophys. Acta* **47**, 612.

Hayashi, M., and Spiegelman, S. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 1564.

Hecht, L., Stephenson, M., and Zamecnik, P. (1959). *Proc. Natl. Acad. Sci. U.S.* **45**, 505.

Herbert, E., and Canellakis, E. (1960). *Biochim. et Biophys. Acta* **42**, 363.

Herbert, E., and Canellakis, E. (1961). *Biochim. et Biophys. Acta* **47**, 85.

Hirokawa, R., Omori, S., Takahashi, T., and Ogata, K. (1961). *Biochim. et Biophys. Acta* **49**, 612.

Hoagland, M. B. (1955). *Biochim. et Biophys. Acta* **16**, 288.

Hoagland, M. B. (1960). *In* "The Nucleic Acids" (E. Chargaff and J. N. Davidson, eds.), Vol. III, pp. 349–408. Academic Press, New York.

Hoagland, M. B., and Comly, L. T. (1960). *Proc. Natl. Acad. Sci. U.S.* **46**, 1554.

Hoagland, M. B., Zamecnik, P. C., and Stephenson, M. L. (1957). *Biochim. et Biophys. Acta* **24**, 215.

Hoagland, M. B., Stephenson, M. L., Scott, J. F., Hecht, L. I., and Zamecnik, P. C. (1958). *J. Biol. Chem.* **231**, 241.

Holley, R. W., and Goldstein, J. (1959). *J. Biol. Chem.* **234**, 1765.

Holley, R., and Merrill, S. (1959). *J. Am. Chem. Soc.* **81**, 753.

Holloway, B., and Ripley, S. (1952). *J. Biol. Chem.* **196**, 695.

Hülsmann, W. C., and Lipmann, F. (1960). *Biochim. et Biophys. Acta* **43**, 123.

Hultin, T. (1950). *Exptl. Cell Research* **1**, 376.

Hultin, T. (1956). *Exptl. Cell Research* **11**, 222.

Hultin, T. (1961). *Biochim. et Biophys. Acta* **51**, 219.

Hurwitz, J., Bresler, A., and Diringer, R. (1960). *Biochem. Biophys. Research Communs.* **3**, 15.

Jacob, F., and Monod, J. (1961). *J. Mol. Biol.* **3**, 318.

Jonxis, J. H. P. (1958). *Exposés ann. biochim. méd.* **20**, 89.

Kameyama, T., and Novelli, G. D. (1960). *Biochem. Biophys. Research Communs.* **2**, 393.

Kihara, H., Hu, A., and Halvorson, H. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 489.

Kruh, J., Dreyfus, J., Schapira, G., and Padieu, P. (1957). *J. Biol. Chem.* **228**, 113.

Kruh, J., Dreyfus, J., Schapira, G. (1960). *J. Biol. Chem.* **235**, 1075.

Kruh, J., Rosa, J., Dreyfus, J., and Schapira, G. (1961). *Biochim. et Biophys. Acta* **49**, 509.

Lacks, S., and Gros, F. (1959). *J. Mol. Biol.* **1**, 301.

Lamborg, M., and Zamecnik, P. (1960). *Biochim. et Biophys. Acta* **42**, 206.

Lamfrom, H. (1961). *J. Mol. Biol.* **3**, 241.

Leahy, J., Glassman, E., and Schweet, R. (1960). *J. Biol. Chem.* **235**, 3209.

Lengyel, P., Speyer, J., and Ochoa, S. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 1936.

Lipmann, F., Hülsmann, W. C., Hartmann, G., Boman, H., and Acs, G. (1959). *J. Cellular Comp. Physiol.* **54** (Suppl. 1), 75.

Loftfield, R. (1960). *Fourth Intern. Congr. Biochem., Vienna* **8**, 222.

McFall, E. (1961). *J. Mol. Biol.* **3**, 219.

McFall, E., Pardee, A. B., and Stent, G. S. (1958). *Biochim. et Biophys. Acta* **27**, 282.

Matthaei, J. H., and Nirenberg, M. W. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 1580.

Matthaei, J. H., Jones, O., Martin, R., and Nirenberg, M. (1962). *Proc. Natl. Acad. Sci. U.S.* **48**, 666.

Miller, A., and Lamfrom, H. (1962). *Federation Proc.* **21**, 413.

Moldave, K., and Grossi, L. G. (1960). *Federation Proc.* **19**, 351.

Morris, A. J., and Dickman, S. (1960). *J. Biol. Chem.* **235**, 1404.

Morris, A. J., and Schweet, R. S. (1961). *Biochim. et Biophys. Acta* **47**, 415.

Morris, A. J., Favelukes, S., Arlinghaus, R., and Schweet, R. (1962). *Biochem. Biophys. Research Communs.* **7**, 326.

Muir, H. M., Neuberger, A., and Perrone, J. C. (1952). *Biochem. J.* **52**, 87.

Naono, S., and Gros, F. (1960a). *Compt. rend. acad. sci.* **250**, 8527.

Naono, S., and Gros, F. (1960b). *Compt. rend. acad. sci.* **250**, 3889.

Nathans, D., and Hülsman, W. C. (1960). *Federation Proc.* **19**, 347.

Nathans, D., and Lipmann, F. (1960). *Biochim. et Biophys. Acta* **43**, 126.

Nathans, D., and Lipmann, F. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 497.

Nathans, D., von Ehrenstein, G., Monroe, R., and Lipmann, F. (1962). *Federation Proc.* **21**, 127.

Nirenberg, M. W., and Matthaei, J. H. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 1588.

Nirenberg, M. W., Matthaei, H., and Jones, O. (1962). *Proc. Natl. Acad. Sci. U.S.* **48**, 104.

Nomura, M., Hall, B. D., and Spiegelman, S. (1960). *J. Mol. Biol.* **2**, 306.

Novelli, G. D., Kameyama, T., and Eisenstadt, J. M. (1962). *Cold Spring Harbor Symposia Quant. Biol.* **26**, 133.

Ogata, K., and Nohara, H. (1957). *Biochim. et Biophys. Acta* **25**, 660.

Ogata, K., Hirokawa, R., and Omori, S. (1960). *Biochim. et Biophys. Acta* **40**, 178.

Osawa, S. (1960). *Biochim. et Biophys. Acta* **43**, 110.

Perry, R. P., Hell, A., and Errera, M. (1961a). *Biochim. et Biophys. Acta* **49**, 47.

Perry, R. P., Hell, A., and Errera, M. (1961b). *Biochim. et Biophys. Acta* **49**, 58.

Peters, T. (1957). *J. Biol. Chem.* **229**, 659.

Pizer, L., and Cohen, S. (1961). *Biochim. et Biophys. Acta* **53**, 409.

Preiss, J., Berg, P., Ofengand, E., Bergmann, F., and Dieckmann, M. (1959). *Proc. Natl. Acad. Sci. U.S.* **45**, 319.

Prescott, D. (1960). *Exptl. Cell Research* **19**, 29.

Raacke, I. D. (1961). *In* "Metabolic Pathways" (D. E. Greenberg, ed.), Vol. II, pp. 263–388. Academic Press, New York.

Rabinowitz, M., and Fisher, J. (1962). *J. Biol. Chem.* **237**, 477.

Rabinowitz, M., and Olson, M. (1959). *J. Biol. Chem.* **234**, 2085.

Rho, J. H., and Bonner, J. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 1611.

Riley, M., Pardee, A. B., Jacob, F., and Monod, J. (1960). *J. Mol. Biol.* **2**, 216.

Roberts, R. (1959). *Carnegie Inst. Wash. Yearbook* **59**, 268.

Schapira, G., Dreyfus, J. C., Kruh, J., and Labie, D. (1960). *Biochim. et Biophys. Acta* **45**, 409.

Schneider, J. H. (1961). *Biochim. et Biophys. Acta* **47**, 107.

Schweet, R., Bovard, F., Allen, E., and Glassman, E. (1958a). *Proc. Natl. Acad. Sci. U.S.* **44**, 173.

Schweet, R., Lamfrom, H., and Allen, E. (1958b). *Proc. Natl. Acad. Sci. U.S.* **44,** 1029.

Schweet, R. S., Bishop, J., and Morris, A. (1961). *Lab. Invest.* **10,** 992.

Shimura, K., Fukai, H., Sato, J., and Saeki, R. (1956). *J. Biochem. (Tokyo)* **43,** 101.

Sibatani, A., de Kloet, S., Allfrey, V., and Mirsky, A. (1962). *Proc. Natl. Acad. Sci. U.S.* **48,** 471.

Siekevitz, P. (1952). *J. Biol. Chem.* **195,** 549.

Siekevitz, P., and Palade, G. (1960). *J. Biophys. Biochem. Cytol.* **7,** 619.

Simkin, J. (1958). *Biochem. J.* **70,** 305.

Singer, M., and Cantoni, G. (1960). *Biochim. et Biophys. Acta* **39,** 182.

Smith, E. (1962). *Proc. Natl. Acad. Sci. U.S.* **48,** 677.

Smith, K., Cordes, E., and Schweet, R. (1959). *Biochim. et Biophys. Acta* **33,** 286.

Speyer, J., Lengyel, P., Basilio, C., and Ochoa, S. (1962). *Proc. Natl. Acad. Sci. U.S.* **48,** 441.

Spiegelman, S., Halvorson, H., and Ben-Ishai, R. (1955). *In* "Amino Acid Metabolism" (W. McElroy and B. Glass, eds.), p. 124. Johns Hopkins Press, Baltimore.

Steinberg, D., Vaughan, M., and Anfinsen, C. B. (1956). *Science* **124,** 389.

Stephenson, M., and Zamecnik, P. (1961). *Proc. Natl. Acad. Sci. U.S.* **47,** 1627.

Stevens, A. (1960). *Biochem. Biophys. Research Communs.* **3,** 92.

Stulberg, M. P., and Novelli, G. D. (1962). *In* "The Enzymes" (P. Boyer and H. Lardy, eds.), Vol. VI. Academic Press, New York. In press.

Suzuka, I., and Shimura, K. (1960). *J. Biochem. (Tokyo)* **47,** 555.

Takanami, M. (1961). *Biochim. et Biophys. Acta* **51,** 85.

Tissières, A., and Hopkins, J. (1961). *Proc. Natl. Acad. Sci. U.S.* **47,** 2015.

Tissières, A., Schlessinger, D., and Gros, F. (1960). *Proc. Natl. Acad. Sci. U.S.* **46,** 1450.

T'so, P., and Sato, C. (1959). *Exptl. Cell Research* **17,** 227.

von der Decken, A., and Hultin, T. (1958). *Exptl. Cell Research* **15,** 254.

von der Decken, A., and Hultin, T. (1960). *Biochim. et Biophys. Acta* **45,** 139.

von Ehrenstein, G., and Lipmann, F. (1961). *Proc. Natl. Acad. Sci. U.S.* **47,** 941.

Webster, G. C. (1960). *Federation Proc.* **19,** 345.

Webster, G. C. (1961). *Federation Proc.* **20,** 387.

Weiss, S. B. (1960). *Proc. Natl. Acad. Sci. U.S.* **46,** 1020.

Weiss, S. B., and Nakamoto, T. (1961). *Proc. Natl. Acad. Sci. U.S.* **47,** 694.

Wood, W., and Berg, P. (1962). *Proc. Natl. Acad. Sci. U.S.* **48,** 94.

Yarmolinsky, M. B., and de la Haba, G. L. (1959). *Proc. Natl. Acad. Sci. U.S.* **45,** 1721.

Yoshida, A., and Tobita, T. (1960). *Biochim. et Biophys. Acta* **37,** 513.

Zachau, H., Acs, G., and Lipmann, F. (1958). *Proc. Natl. Acad. Sci. U.S.* **44,** 885.

Zalokar, M. (1960). *Exptl. Cell Research* **19,** 114.

Zamecnik, P. C. (1960). *Harvey Lectures* **54,** 256.

Zamecnik, P., and Keller, E. (1954). *J. Biol. Chem.* **209,** 337.

Zillig, W., Schachtschabel, D., and Krone, W. (1960). *Z. physiol. Chem. Hoppe-Seyler's* **318,** 100.

## *Chapter* IX

# Correlations between Genetics and Chemistry of Human Hemoglobins

CORRADO BAGLIONI[1]

*Department of Biology,*
*Massachusetts Institute of Technology,*
*Cambridge, Massachusetts*

I. Introduction . . . . . . . . . . . . . . . . . . 405
II. The Separation of Human Hemoglobins by Chromatography and Electrophoresis . . . . . . . . . . . . . . . . . 406
III. Normal Hemoglobins . . . . . . . . . . . . . . . . 409
  A. Hb-A . . . . . . . . . . . . . . . . . . . 409
  B. Hb-F . . . . . . . . . . . . . . . . . . . 415
  C. Hb-$A_2$ . . . . . . . . . . . . . . . . . . 418
  D. Considerations on the Amino Acid Sequence of the Human Hemoglobin Chains . . . . . . . . . . . . . . 419
  E. Hb-$A_3$ . . . . . . . . . . . . . . . . . . 420
  F. Embryonic Hemoglobin . . . . . . . . . . . . . 421
IV. Abnormal Hemoglobins . . . . . . . . . . . . . . 421
  A. Hb-A Variants—Chemical Studies . . . . . . . . . 423
  B. Formal Genetics of the Hb-A Variants . . . . . . . 436
  C. Genetic and Biochemical Relationships . . . . . . . 438
  D. Dissociation and Recombination of Hemoglobin . . . . 440
  E. Hb-F Variants . . . . . . . . . . . . . . . 442
  F. Hb-$A_2$ Variants . . . . . . . . . . . . . . 444
  G. The Problem of Thalassemia . . . . . . . . . . 449
  H. Abnormal Hemoglobins from Altered Aggregation of Peptide Chains . . . . . . . . . . . . . . . 453
V. The Control of the Synthesis of Hemoglobin . . . . . . . 456
  A. The Control of the Rate of Synthesis of the Peptide Chains . . . 456
  B. The Control of the Synthesis of Fetal and Adult Hemoglobin . . 459
  C. Hereditary Persistence of Hb-F . . . . . . . . . 462
  D. Models of Red Cell Differentiation and Synthesis of Hb-F . . . 466
  Acknowledgments . . . . . . . . . . . . . . . . 469
  References . . . . . . . . . . . . . . . . . 469

## I. Introduction

Substantial advances in our knowledge of the genetics and chemistry of the human hemoglobins have been made recently and the attention of several investigators has been driven to this field of research. A rough estimate of the number of papers published within the last ten years on this subject shows this number to be close to a thousand. It is impossible to review and give reference to all these papers for reasons of conciseness and limitation of space. The necessary omissions should not imply lack of appreciation of the papers not cited; most of these references have not been included because they are reported in previously published reviews.

---

[1] Present address: International Laboratory for Genetics and Biophysics, Naples, Italy.

The rapid progress in the past few years has stimulated many authors to cover this field of research with excellent and specialized reviews. Among the many good reviews may be mentioned Itano's (1957) review on human hemoglobins, Lehmann's (1960) review on abnormal hemoglobins, Ingram's (1961b) book "Hemoglobin and Its Abnormalities," and Rucknagel and Neel's (1961) review on hemoglobinopathies. Gratzer and Allison (1960) have recently reviewed the related field of the animal hemoglobins. Most of these reviews report the historical developments of the research in the particular field covered, so that no attempt will be made to present a historical background in this review.

The purpose of this review is to give the reader up-to-date information on the latest developments of the genetic and chemical studies on the human hemoglobins and to correlate the work of people active in widely different disciplines: human genetics, biochemistry, hematology, and biophysics.

## II. The Separation of Human Hemoglobins by Chromatography and Electrophoresis

Hemoglobin is a conjugated protein, consisting of a protein part (globin) and of a prosthetic group (heme). Mammalian hemoglobins are rather large molecules, made up of approximately 600 amino acids and with a molecular weight of 66,000 ± 2,000. Hemoglobin is contained in specialized cells, the red cells, which are nucleated in lower vertebrates and anucleated in mammals. Red cells are easily obtained by bleeding and hemoglobin solutions are usually prepared by lysis of the red cells with hypotonic solutions, after washing away the serum proteins with isotonic solutions. Over 95% of the soluble protein obtained in this way from mammalian red cells is hemoglobin. This remarkable circumstance and the ready availability of large quantities of red cells have made hemoglobin one of the preferred objects of study for the protein chemist for almost a century.

Hemoglobins obtained from different animal species have been analyzed by several investigators; the heme has been found to be identical in all hemoglobins (protoporphyrin IX + Fe), while the globin varies considerably in physicochemical properties and in amino acid composition in different species. In many cases more than one hemoglobin has been demonstrated in individuals of the same species by the electrophoretic and chromatographic methods devised for protein separation. That heterogeneity of hemoglobin arises as a consequence of the oxidation or reduction of the heme, can usually be excluded by converting the heme to the carbomonoxy or cyanmet derivative, both of which have the same electrophoretic charge as oxyhemoglobin. Consequently, the hemoglobin found in individuals and in species with multiple hemoglobins differ somewhat in the protein part of the molecule.

Pauling *et al.* discovered in 1949 the difference in electrophoretic mobility between normal human hemoglobin and hemoglobin obtained from sickle-cell anemic patients. This observation has been of fundamental importance in the development of the concept of gene-protein relationship and has stimulated a number of workers to investigate human hemoglobins. Several abnormal hemoglobins have been discovered by the same method of electrophoretic analysis. The original Tiselius moving boundary electrophoresis technique was found, however, to be an expensive and relatively impractical method, because of the amount of work involved in a single determination and because of the expensive equipment. Simpler methods of analysis (such as paper or starch-gel electrophoresis) have been developed for the electrophoretic analysis of the human hemoglobins. These methods have been applied to the study of "normal" human hemoglobin and of the human hemoglobin variants.

"Normal" human hemoglobin is not homogeneous in free boundary electrophoresis (Hoch, 1950). Kunkel and Wallenius (1955) isolated a major and two minor components by using starch block electrophoresis. This method of electrophoretic analysis allows the elution from the starch slab of the separated hemoglobins and the quantitative evaluation of each component.

The more sensitive starch-gel electrophoresis (Smithies, 1959) has now become more widely used in order to reveal minor components; the buffer system described by Poulik (1957) exhibits considerable sharpening of the bands and is greatly advantageous. The sliced starch-gel is stained with the benzidine or with the *o*-dianisidine reagent (Smithies, 1959), which are catalytically oxidized by hemoproteins in the presence of hydrogen peroxide. The sensitivity of the method is considerably enhanced and components present in quantities of less than 1% of the total hemoglobin can be detected. The use of filter paper electrophoresis is confined to the screening of major abnormal components. Adsorption on the filter paper and tailing of the hemoglobins may preclude the observation of minor components having an electrophoretic mobility lower than the major components. Filter paper electrophoresis is still the simplest and fastest method of analysis for the detection of major components if a large number of samples have to be analyzed. The electrophoretic analysis on agar gels has also been applied to the separation of hemoglobins (see review by Gratzer and Beaven, 1961).

Chromatographic methods of analysis of human hemoglobin on ion exchangers have also been developed; these methods offer better resolution and quantitative isolation of the separated components. The presence of minor components in samples of human hemoglobin has been confirmed by several investigators, using different analytical techniques for protein separation. Huisman and Meyering (1960) have standardized the chromatographic separation on carboxymethylcellulose, described by Peterson and Sober (1956), for the qualitative analysis of human hemoglobin and for the preparative isolation of the components.

Morrison and Cook (1955) demonstrated by chromatography on the resin IRC-50 that normal adult hemoglobin is heterogeneous; Prins and Huisman (1956) also detected heterogeneity of human hemoglobin in the same way. The chromatographic method of analysis on IRC-50 has been greatly developed and improved by Allen *et al.* (1958) and subsequently by Clegg and Schroeder (1959) and by Schnek and Schroeder (1961); several minor components are separated by this method of analysis.

The separation of human hemoglobins by starch block electrophoresis and by column chromatography has considerably complicated the problem of the identification of components isolated by different techniques and the problem of nomenclature.

The original system of nomenclature proposed by Kunkel and Wallenius (1955) has been the most widely accepted. The major component isolated by starch block electrophoresis has been designated $A_1$ by these authors; it represents approximately 90% of the total hemoglobin. The most slowly moving hemoglobin component in electrophoresis has been designated $A_2$ and the most rapidly moving $A_3$; they represent 2–3% and 4–12% of the total hemoglobin, respectively. It has become customary to indicate the major component as hemoglobin A or in shortened form as Hb-A. However, the two minor components are still indicated as Hb-$A_2$ and Hb-$A_3$. Keys to the identification of the minor components isolated by starch block electrophoresis with the components isolated by column chromatography are found in the articles of Schnek and Schroeder (1961) and of Huisman and Meyering (1960).

## III. Normal Hemoglobins

### A. Hb-A

Hb-A is the normal major component of hemoglobin prepared from the red cells of adult individuals. Hb-A is usually defined by its electrophoretic mobility, the isoelectric point being at pH 6.87, and by its rate of alkali and acid denaturation. Most of the physical and chemical properties established by the protein chemists in the past for human hemoglobin can be referred to Hb-A. Rhinesmith *et al.* (1957a) determined the N-terminal residues of Hb-A; four *valine* residues were obtained indicating the presence of four peptide chains. Rhinesmith *et al.* (1957b) next found that two peptide chains on acid hydrolysis release quickly the N-terminal dipeptide *valyl-leucine* and two chains release *valine* only more slowly; these peptide chains have been designated $\alpha$ chains and $\beta$ chains, respectively. The N-terminal sequence of the latter has been found to be Val.His.Leu. (Rhinesmith *et al.*, 1958).[2]

The presence of two types of peptide chains in Hb-A, demonstrated by Rhinesmith *et al.* (1958), has been of particular importance in solving some of the genetic and chemical problems presented by the abnormal hemoglobins. The idea of a molecule composed of two pairs of identical chains agreed particularly well with the X-ray crystallographic picture of the horse hemoglobin molecule, obtained by Perutz and his collaborators (see Perutz *et al.*, 1960). The horse hemoglobin molecule has the shape of an ellipsoid composed of two identical parts. In each half-molecule there are two different peptide chains. This type of structure is probably common to all mammalian hemoglobins. A rather schematic representation of the hemoglobin molecule is shown in Fig. 1.

Wilson and Smith (1959) obtained pure preparations of the two peptide chains of horse hemoglobin by ion exchange chromatography in a urea gradient and by preparative boundary electrophoresis. Ingram (1959a) applied the same methods to the preparation of the peptide chains of human adult hemoglobin; two fractions were separated by free boundary electrophoresis and by column chromatography.

The fractions isolated by column chromatography were examined by "fingerprinting." In the fingerprinting analysis a mixture of peptides, obtained by tryptic hydrolysis of denatured hemoglobin, is separated by paper ionophoresis, followed by paper chromatography in a perpendicular direction. The peptides are spread on paper sheets in character-

---

[2] Amino acids are indicated by the first three letters of their names. Asparagine and glutamine are indicated as Asp-$NH_2$ and Glu-$NH_2$, respectively, in the text, but as Asn and Gln in the illustrations. Isoleucine is indicated as Ileu.

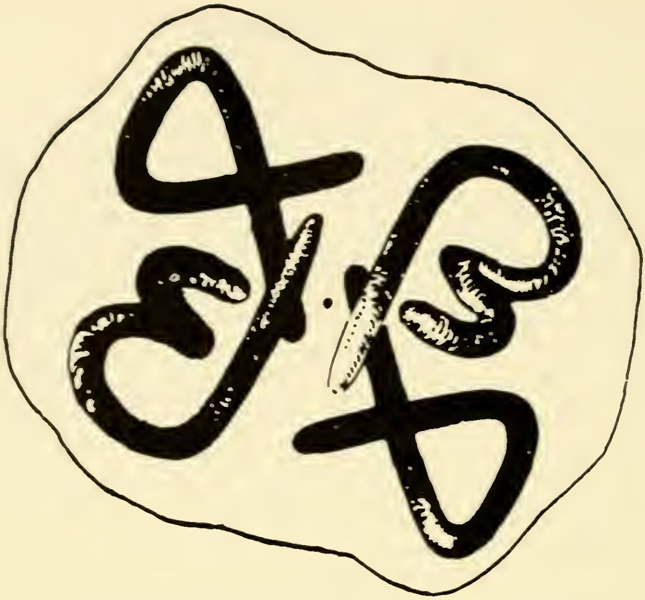


FIG. 1. Schematic representation of the arrangement of the peptide chains in the hemoglobin molecule. (From Ingram and Stretton, 1959b; reproduced by permission of the editor of *Nature*.)

istic patterns and the comparison of patterns of peptides—fingerprints—allows the detection of minor differences in the composition of the proteins under analysis (see Fig. 2, A and B).

The fingerprinting method has been modified by several authors, who have adapted the same principle, i.e., the two-dimensional separation by ionophoresis and chromatography, to different types of ionophoresis apparatuses or to different chromatographic systems. These methods of analysis yield comparable results; in the author's experience the original (Ingram's) method (1958) is to be recommended for its simplicity and practicality. The use of a solvent system different from the one originally employed by Ingram (1958) has been found advantageous (Baglioni, 1961).

The fingerprints of the isolated chains of Hb-A showed different patterns of peptides, which taken together made up the pattern given by whole hemoglobin (Ingram, 1958, 1959a). The fractions were then identified with the two types of peptide chains described by Rhinesmith *et al.* (1958), by determining their N-terminal sequence (Ingram, 1959a). A great deal about the primary sequence of Hb-A has been learned recently and the amino acid sequence of the two chains of Hb-A is now completely elucidated.

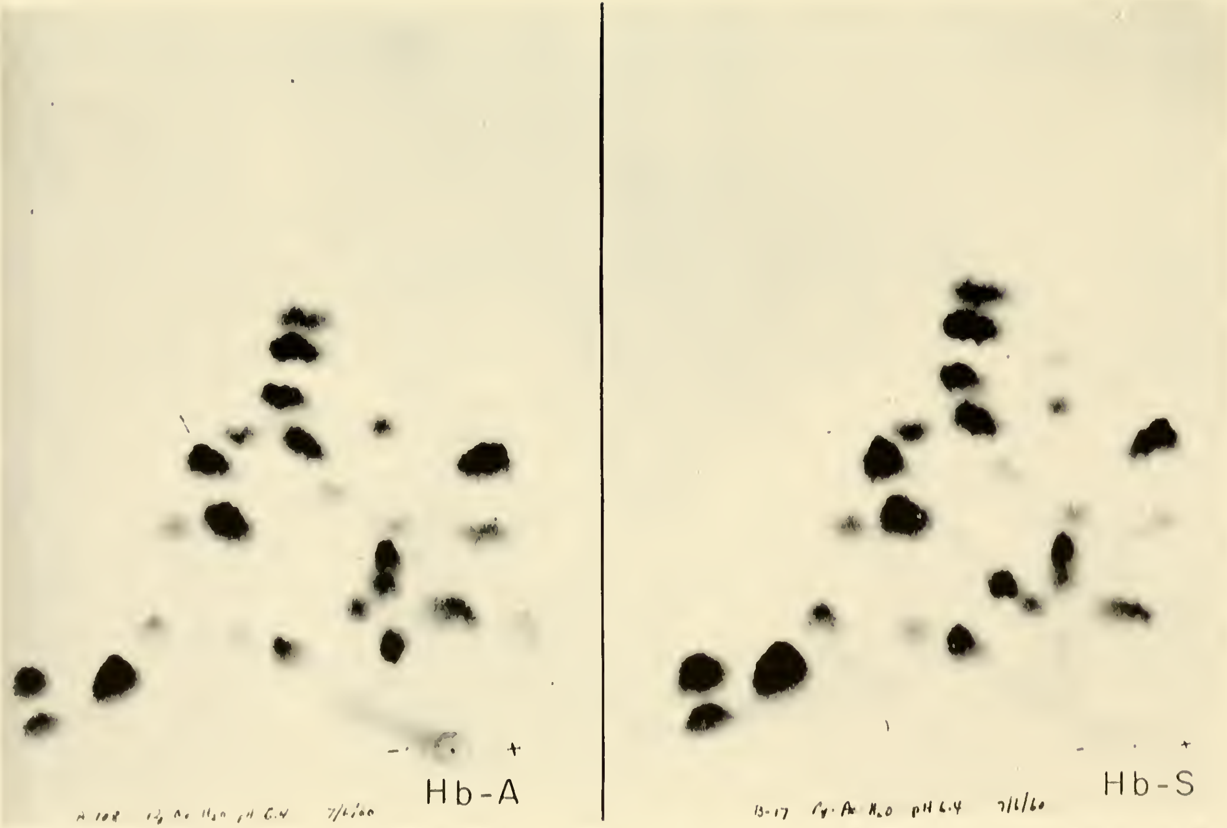


FIG. 2A. Fingerprints of Hb-A and Hb-S (Baglioni, 1961). Photographs of the fingerprints. See also Fig. 2B, p. 412.

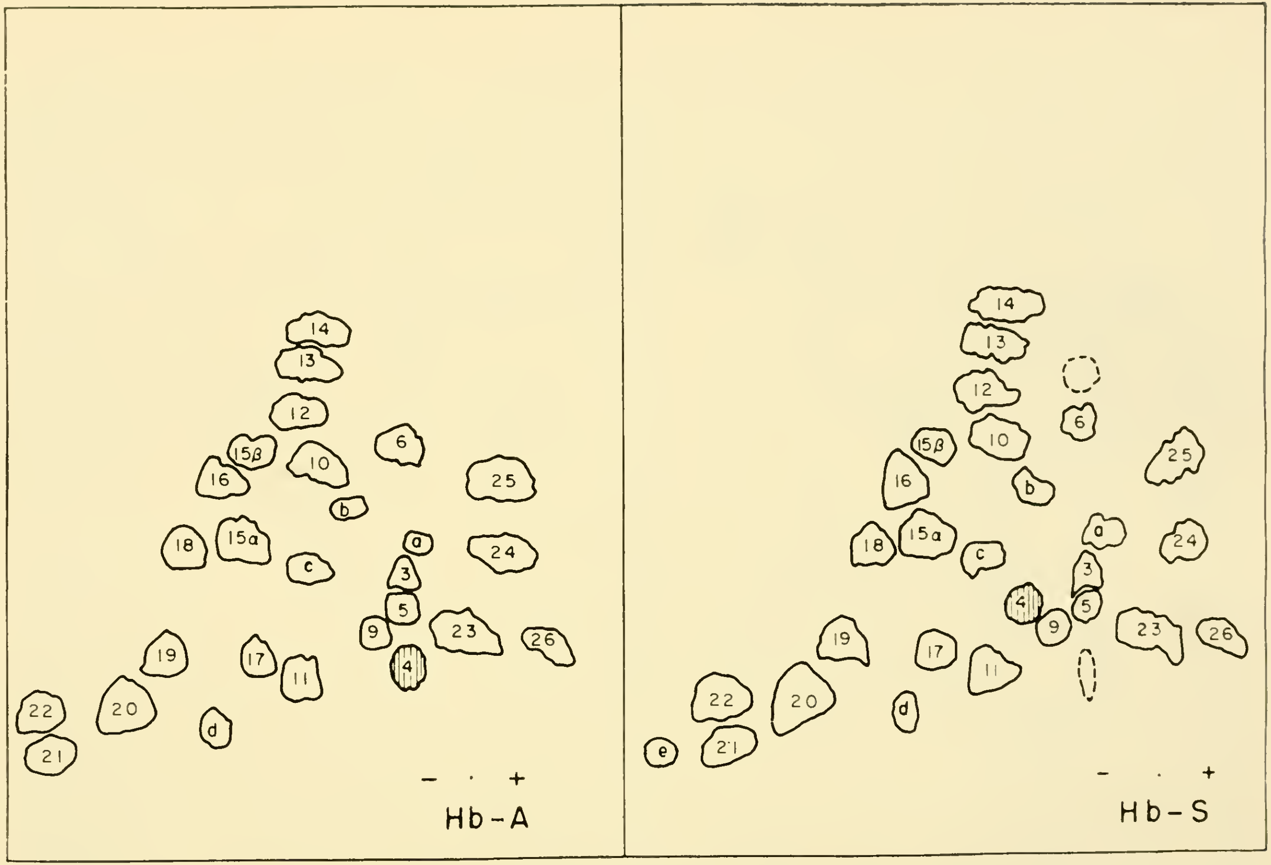

14
13
12
15β
10
6
16
25
b
18
15a
a
c
24
3
5
9
23
26
19
17
11
4
22
20
d
21
− · +
Hb-A
14
13
12
6
15β
10
16
25
b
a
18
15a
c
24
3
4
5
9
23
26
19
17
11
22
20
d
e
21
− · +
Hb-S

FIG. 2B. Tracings of the fingerprints of Hb-A and Hb-S (Baglioni, 1961). Electrophoresis at pH 6.4 for 2½ hours in the horizontal direction; chromatography in isoamyl alcohol-pyridine-water (35:35:30) in the vertical direction. For identification of the tryptic peptides and explanations see Table I. (Reproduced by permission of the Elsevier Publishing Company.)

It is beyond the purpose of this review to give a full account of the very specialized chemical methods that have been used to study the amino acid sequence of the hemoglobin chains. The peptide chains have been separated by column chromatography, according to Wilson and Smith (1959), or by countercurrent distribution (Hill and Craig, 1959). Hydrolysis by means of proteolytic enzymes has been used to degrade the peptide chains to small peptides. Trypsin was the first choice because trypsin specifically cleaves the peptide bonds between the carboxyl group of lysine or arginine and the amino group of other amino acids. The separation and purification of the tryptic peptides (peptides obtained

TABLE I

TRYPTIC PEPTIDES OF THE $\alpha$ AND $\beta$ CHAINS OF Hb-A

| $\alpha$ chain peptides | | $\beta$ chain peptides | |
|---|---|---|---|
| Peptide number[a] | Peptide position[b] | Peptide number[a] | Peptide position[b] |
| 3 | $\alpha$Tp IX | 4 | $\beta$Tp I |
| 9 | $\alpha$Tp VIII, IX | 5 | $\beta$Tp XIII |
| 10 | $\alpha$Tp VI | 6 | $\beta$Tp IX |
| 11 | $\alpha$Tp I, II | 12 | $\beta$Tp II |
| 13 | $\alpha$Tp V | 14 | $\beta$Tp IV |
| 15$\alpha$ | $\alpha$Tp III | 15$\beta$ | $\beta$Tp XV |
| 16$\alpha$ | $\alpha$Tp XIV | 16$\beta$ | $\beta$Tp XIV |
| 17$\alpha$ | $\alpha$Tp II | 17$\beta$ | $\beta$Tp XIV, XV |
| 18 | $\alpha$Tp X | 19 | $\beta$Tp VI |
| 20$\alpha$ | $\alpha$Tp VII | 20$\beta$ | $\beta$Tp VII |
| 21$\alpha$ | $\alpha$Tp VII, VIII | 21$\beta$ | $\beta$Tp VII, VIII |
| 22$\alpha$ | $\alpha$Tp VIII | 22$\beta$ | $\beta$Tp VIII |
| 23 | $\alpha$Tp IV | 24 | $\beta$Tp V oxidized[c] |
| c | $\alpha$Tp I | 25 | $\beta$Tp V |
| | | 26 | $\beta$Tp III |
| | | a | $\beta$Tp II, III |
| | | b | $\beta$Tp VIII, IX |

[a] Peptide numbers refer to Fig. 2B, according to Ingram (1958). Two peptides are present in each of the spots 16, 20, and 21, one from the $\alpha$ chain and one from the $\beta$ chain. Peptide 17$\beta$ overlaps with peptide 15$\alpha$ in Fig. 2.

[b] The tryptic peptides are numbered consecutively from the N-terminus in each chain (Gerald and Ingram, 1961). Peptide 22 happens to be free lysine. The identification of the tryptic peptides and the correspondence with the amino acid sequence of the peptide chains (Braunitzer *et al.*, 1961b; Konigsberg *et al.*, 1961) are based on the amino acid composition and/or sequence of the peptides (quoted by Baglioni, 1961). Preliminary evidence exists for the identification of peptides a, b, and c; analysis of peptide e (Fig. 2) gives only lysine (C. Baglioni, unpublished). Peptide e may possibly be lysyl-lysine.

[c] Peptide 24 is identical with peptide 25, except that the methionine residue in this peptide has become oxidized.

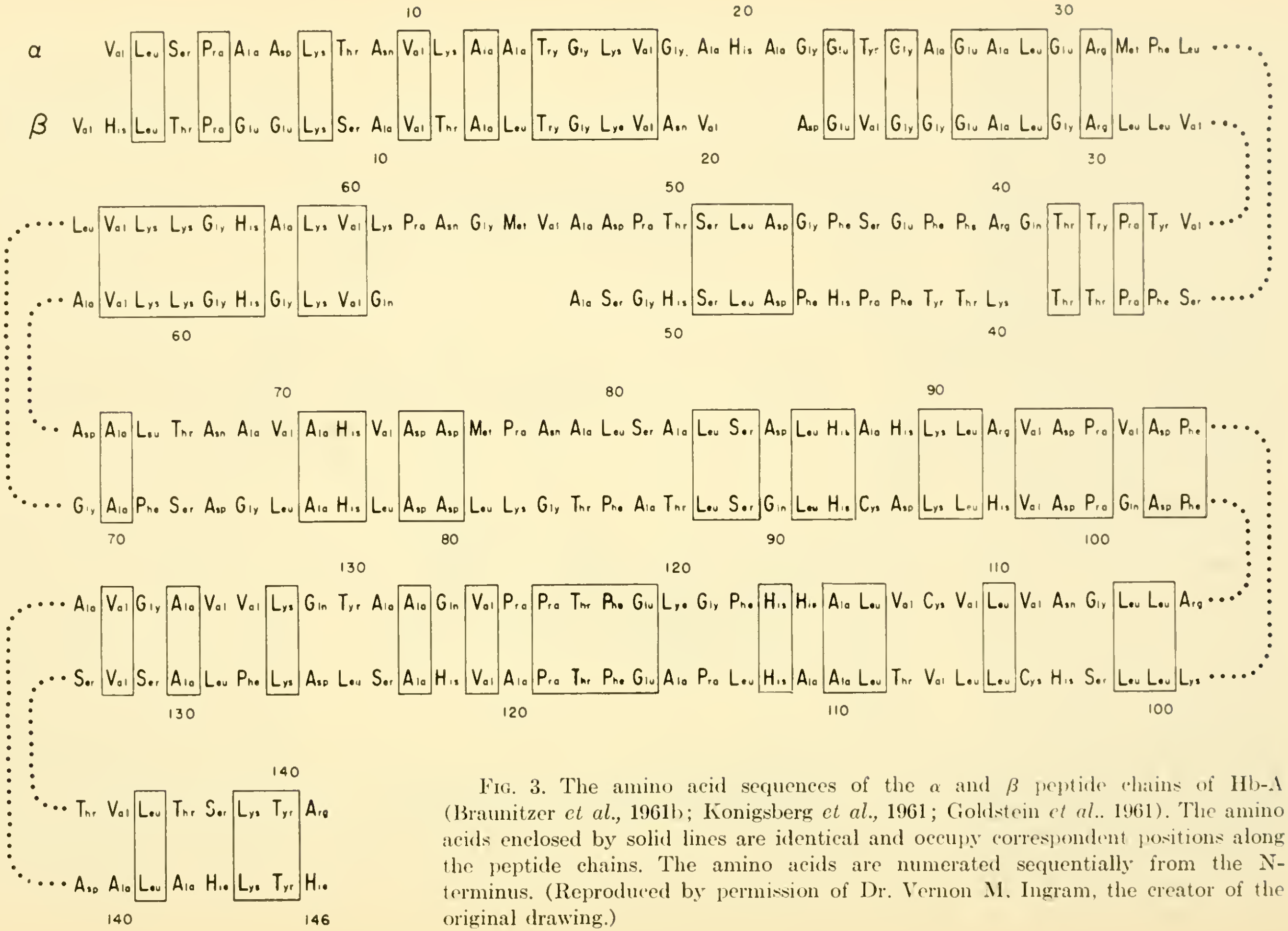

α
β
10
20
30
Val Leu Ser Pro Ala Asp Lys Thr Asn Val Lys Ala Ala Try Gly Lys Val Gly. Ala His Ala Gly Glu Tyr Gly Ala Glu Ala Leu Glu Arg Met Phe Leu
Val His Leu Thr Pro Glu Glu Lys Ser Ala Val Thr Ala Leu Try Gly Lys Val Asn Val Asp Glu Val Gly Gly Glu Ala Leu Gly Arg Leu Leu Val
10
20
30
60
50
40
Leu Val Lys Lys Gly His Ala Lys Val Lys Pro Asn Gly Met Val Ala Asp Pro Thr Ser Leu Asp Gly Phe Ser Glu Phe Phe Arg Gln Thr Try Pro Tyr Val
Ala Val Lys Lys Gly His Gly Lys Val Gln Ala Ser Gly His Ser Leu Asp Phe His Pro Phe Tyr Thr Lys Thr Thr Pro Phe Ser
60
50
40
70
80
90
Asp Ala Leu Thr Asn Ala Val Ala His Val Asp Asp Met Pro Asn Ala Leu Ser Ala Leu Ser Asp Leu His Ala His Lys Leu Arg Val Asp Pro Val Asp Phe
Gly Ala Phe Ser Asp Gly Leu Ala His Leu Asp Asp Leu Lys Gly Thr Phe Ala Thr Leu Ser Gln Leu His Cys Asp Lys Leu His Val Asp Pro Gln Asp Phe
70
80
90
100
130
120
110
Ala Val Gly Ala Val Val Lys Gln Tyr Ala Ala Gln Val Pro Pro Thr Phe Glu Lys Gly Phe His His Ala Leu Val Cys Val Leu Val Asn Gly Leu Leu Arg
Ser Val Ser Ala Leu Phe Lys Asp Leu Ser Ala His Val Ala Pro Thr Phe Glu Ala Pro Leu His Ala Ala Leu Thr Val Leu Leu Cys His Ser Leu Leu Lys
130
120
110
100
140
Thr Val Leu Thr Ser Lys Tyr Arg
Asp Ala Leu Ala His Lys Tyr His
140
146

FIG. 3. The amino acid sequences of the α and β peptide chains of Hb-A (Braunitzer *et al.*, 1961b; Konigsberg *et al.*, 1961; Goldstein *et al.*, 1961). The amino acids enclosed by solid lines are identical and occupy correspondent positions along the peptide chains. The amino acids are numerated sequentially from the N-terminus. (Reproduced by permission of Dr. Vernon M. Ingram, the creator of the original drawing.)

by enzymatic digestion with trypsin) is a troublesome problem and present several difficulties. Liebold and Braunitzer (1959) have used column chromatography, followed by paper ionophoresis and/or paper chromatography. Hill and Konigsberg (1960) have used countercurrent distribution, followed by column chromatography and paper ionophoresis.

The composition of the tryptic peptides has been determined by Braunitzer *et al.* (1960) and by Hill and Konigsberg (1960). The hemoglobin chains were digested by means of proteolytic enzymes characterized with different specificities, like chymotrypsin or pepsin, so that overlapping peptides were obtained; i.e., peptides containing either two or more tryptic peptides or parts of two tryptic peptides and bridging them together, so that their sequence could be determined (Braunitzer *et al.*, 1961; Hill and Konigsberg, 1961). The position along the peptide chains of Hb-A of the peptides separated by fingerprinting is indicated in Table I.

By the conventional methods of the protein chemists, the partial sequences obtained have been filled with all the amino acids in definite order. Conclusive results have been published for the $\alpha$ chain of Hb-A by Konigsberg *et al.* (1961), for the $\alpha$ and $\beta$ chain by Braunitzer *et al.* (1961b), and for part of the $\beta$ chain by Goldstein *et al.* (1961). Many years of work of several investigators are summarized in these few lines. This work should be fully appreciated; it has required the constant application of the ingenuity and skill of several experienced investigators and has been concluded in a surprisingly short time. The two sequences of the $\alpha$ chain, which have been independently determined, show complete agreement. The amino acid sequences of the Hb-A peptide chains are shown in Fig. 3.

### B. Hb-F

The major component of hemoglobin prepared from fetal blood is fetal hemoglobin, indicated as Hb-F. Hb-F usually accounts for 70–80% of the total hemoglobin of a full-term infant (Jonxis, 1959); adult hemoglobin appears in the red cells of fetuses at about the thirteenth week of fetal life (Walker and Turnbull, 1955) and increases in relative proportion until it completely replaces Hb-F by the end of the first year of postnatal life. Hb-F is, however, observed in adults affected by acquired or hereditary anemias, like thalassemia major or sickle-cell anemia. High levels of Hb-F have also been found in some healthy Negro adults (Edington and Lehmann, 1955); this condition has been designated "hereditary persistence of Hb-F" (Jacob and Raper, 1958).

Hb-F is characterized by an electrophoretic mobility only slightly different from that of Hb-A. Hb-F is usually characterized and measured by alkali denaturation. The higher resistance to alkali is in fact the outstanding characteristic of Hb-F. The reaction which hemoglobin undergoes in alkali is essentially the change of hemoglobin to hemochromogen (Drabkin and Austin, 1935); the rate of this change can be followed quantitatively by spectrophotometry. The methods for the identification and evaluation of Hb-F have been reviewed by White and Beaven (1959).

The end group analysis of the Hb-F peptide chains has shown the presence of two pairs of peptide chains with the N-terminal sequence Val. Leu. and Gly.His.Phe., respectively (Schroeder and Matsuda, 1958). The peptide chains Gly.His.Phe., which were found to be different from the $\beta$ chains of Hb-A on the basis of their N-terminal sequence, were designated $\gamma$ chains.

Hunt (1959) has isolated by column chromatography the peptide chains of Hb-F and has examined the pattern of peptides after tryptic hydrolysis. The $\alpha$ chains isolated from Hb-A or Hb-F gave identical fingerprints; the fingerprints of the $\beta$ and $\gamma$ chains were found to be different. The chemical analysis of the N-terminal sequences of Hb-A and Hb-F and the fingerprinting analysis support the claim that these hemoglobins have a pair of peptide chains in common, the $\alpha$ chains, which are associated with pairs of different chains, $\beta$ chains in Hb-A and $\gamma$ chains in Hb-F. A sequence of the $\alpha$ chain of Hb-F has been reported by Schroeder *et al.* (1961); this sequence is in agreement with the sequence of the $\alpha$ chain of Hb-A reported by Konigsberg *et al.* (1961) and by Braunitzer *et al.* (1961b). Schroeder *et al.* (1961, 1962) have determined the sequence of the $\gamma$ chain of Hb-F. This sequence is reported in Fig. 4 together with the sequence of the $\beta$ chain of Hb-A. Genetic evidence showing that the $\alpha$ chains of Hb-F and the $\alpha$ chains of Hb-A are under the control of the same gene has been accumulating in the past few years (Baglioni *et al.*, 1961; Minnich *et al.*, 1962), so that there is little doubt today that these peptide chains are chemically identical (see Section IV,E).

Since Hb-A and Hb-F have in common the same $\alpha$ chains, the higher alkali resistance of Hb-F has been attributed to a property of the $\gamma$ chains. The denaturation of Hb-F at alkaline pH's follows, however, a first-order kinetics, with no break in the linear plot corresponding to the presence of a resistant portion of the molecule (Charlwood *et al.*, 1960). Hb-A is known to dissociate at pH 11.0 into symmetric subunits (see Section IV,D) of molecular weight approximately half that of hemoglobin (Hasserodt and Vinograd, 1959). Above pH 11.0, Hb-A dissociates into subunits of smaller molecular weight, presumably into

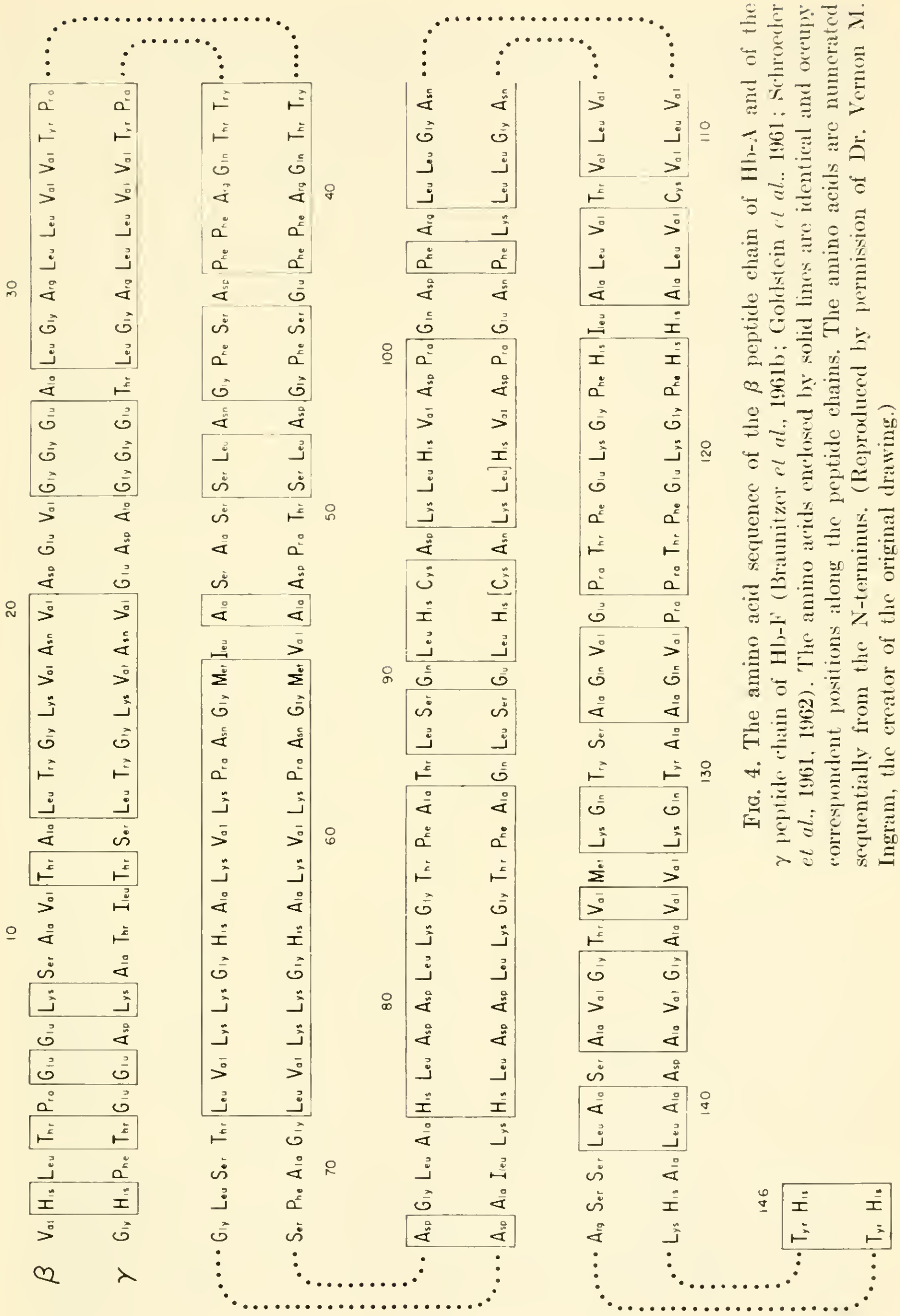

β Val His Leu Thr Pro Glu Glu Lys Ser Ala Val Thr Ala Leu Try Gly Lys Val Asn Val Asp Glu Val Gly Gly Glu Ala Leu Gly Arg Leu Leu Val Val Tyr Pro
γ Gly His Phe Thr Glu Glu Asp Lys Ala Thr Ileu Thr Ser Leu Try Gly Lys Val Asn Val Glu Asp Ala Gly Gly Glu Thr Leu Gly Arg Leu Leu Val Val Tyr Pro
10
20
30
Try Thr Gln Arg Phe Phe Asp Ser Phe Gly Asn Leu Ser Ser Ala Ser Ala Ileu Met Gly Asn Pro Lys Val Lys Ala His Gly Lys Lys Val Leu Thr Ser Leu Gly
Try Thr Gln Arg Phe Phe Glu Ser Phe Gly Asp Leu Ser Thr Pro Asp Ala Val Met Gly Asn Pro Lys Val Lys Ala His Gly Lys Lys Val Leu Gly Ala Phe Ser
40
50
60
70
Asp Gly Leu Ala His Leu Asp Asp Leu Lys Gly Thr Phe Ala Thr Leu Ser Gln Leu His Cys Asp Lys Leu His Val Asp Pro Gln Asp Phe Arg Leu Leu Gly Asn
Asp Ala Ileu Lys His Leu Asp Asp Leu Lys Gly Thr Phe Ala Gln Leu Ser Glu Leu His Cys Asn Lys Leu His Val Asp Pro Glu Asn Phe Lys Leu Leu Gly Asn
80
90
100
Val Leu Val Thr Val Leu Ala Ileu His Phe Gly Lys Glu Phe Thr Pro Glu Val Gln Ala Ser Try Gln Lys Met Val Thr Gly Val Ala Ser Ala Leu Ser Ser Arg
Val Leu Val Cys Val Leu Ala His His Phe Gly Lys Glu Phe Thr Pro Pro Val Gln Ala Ala Tyr Gln Lys Val Val Ala Gly Val Ala Asp Ala Leu Ala His Lys
110
120
130
140
146
Tyr His
Tyr His

FIG. 4. The amino acid sequence of the β peptide chain of Hb-A and of the γ peptide chain of Hb-F (Braunitzer *et al.*, 1961b; Goldstein *et al.*, 1961; Schroeder *et al.*, 1961, 1962). The amino acids enclosed by solid lines are identical and occupy correspondent positions along the peptide chains. The amino acids are numerated sequentially from the N-terminus. (Reproduced by permission of Dr. Vernon M. Ingram, the creator of the original drawing.)

single chains; the changes in the Hb-A molecules exposed to pH 11.6 are not reversible and extensive denaturation occurs (Hasserodt and Vinograd, 1959). The dissociation of Hb-F into half-molecules occurs at higher pH's compared to Hb-A; the dissociation is completed at pH 11.6 and the Hb-F retains at this pH the spectroscopic characteristics of undenatured hemoglobin for some days (Charlwood *et al.*, 1960). Since denaturation of Hb-A occurs at pH's where the molecule appears to be dissociated into single peptide chains, it may be supposed that Hb-F is more resistant to alkali than Hb-A, because of the dissociation equilibrium of Hb-F into single chains being shifted toward higher pH's compared to Hb-A.

The nature of the chemical bonds involved in stabilizing the hemoglobin molecules is not known; it has been suggested that histidines, lysines, and tyrosines provide the active groups involved (Hasserodt and Vinograd, 1959). The different alkali resistance of Hb-A and Hb-F may be due to different chemical bonds established between the peptide chains of these hemoglobins. The higher alkali resistance of Hb-F may not be peculiar of the $\gamma$ chains, but may rather result from the chemical bonds established between the $\alpha$ and $\gamma$ chains in Hb-F being different from those established between the $\alpha$ and $\beta$ chains in Hb-A. A number of factors are involved in determining the higher alkali resistance of Hb-F; aging of the Hb-F, oxidation of the heme group (Matsuda *et al.*, 1960), and exposure to pH 5.0 (Huisman, 1961) are effective in lowering the alkali resistance of Hb-F. Changes of the Hb-F molecule presumably occur under these conditions, which alter the bonds involved in stabilizing the quaternary structure of the molecule.

### C. Hb-$A_2$

This minor component was first described and isolated by Kunkel and Wallenius (1955); it represents approximately 2.5% of the hemoglobin of normal adults. Hb-$A_2$ is easily separated from Hb-A since it is characterized by an electrophoretic mobility at pH 8.6 considerably lower than that of Hb-A. Hb-A and Hb-$A_2$ have identical molecular weights and UV spectra (Kunkel and Bearn, 1957) and very similar amino acid composition (Ingram and Stretton, 1961). Hb-$A_2$ is not a chemical derivative of Hb-A (Kunkel and Bearn, 1957).

Kunkel *et al.* (1957) reported that the Hb-$A_2$ level is approximately doubled in thalassemia minor, the heterozygous, nonsymptomatic form of the hereditary anemia known as thalassemia major or Cooley's anemia. This observation has been confirmed in a larger number of individuals by other investigators (Ceppellini, 1959a; Silverstroni *et al.*, 1957).

Ingram and Stretton (1959a) investigated the chemical nature of Hb-$A_2$ by fingerprinting and determined the sequence of a few peptides of this hemoglobin. Few differences were observed in the fingerprint of Hb-$A_2$, when compared to the fingerprint of Hb-A. The differences were localized to $\beta$ chain peptides. Muller and Jonxis (1960) and Ingram and Stretton (1961) separated the peptide chains of Hb-$A_2$ and fingerprinted the isolated chains. Hb-$A_2$ appears to have a pair of chains identical to the $\alpha$ chains of Hb-A; peptide chains different from the $\beta$ chains of Hb-A are associated with the $\alpha$ chains in Hb-$A_2$; these peptide chains have been designated $\delta$ chains. Ingram and Stretton (1961, 1962) isolated several of the peptides of the $\delta$ chain of Hb-$A_2$ and determined their amino acid composition and some of their amino acid sequences. The $\delta$ chains of Hb-$A_2$ have been found to be surprisingly similar to the $\beta$ chains of Hb-A. Only 4 residues out of 146 have been shown to be definitely different between these two chains (Ingram and Stretton, 1961) and there is preliminary evidence for at least 4 further changes (A. O. W. Stretton, personal communication).

## D. CONSIDERATIONS ON THE AMINO ACID SEQUENCE OF THE HUMAN HEMOGLOBIN CHAINS

In Figs. 3 and 4 the amino acid sequences of the $\alpha$, $\beta$, and $\gamma$ chains are shown together in order to point out the similarities existing between these peptide chains. In the case of the $\beta$ and $\gamma$ chains the similarities are striking. If one calculates the relative number of amino acids in corresponding positions along these peptide chains which are found to be identical, one can observe 42% correspondence between the $\alpha$ and the $\beta$ chains, 39% correspondence between the $\alpha$ and the $\gamma$ chains, and 71% correspondence between the $\beta$ and the $\gamma$ chains. The $\gamma$ chain is certainly more similar to the $\beta$ chain than to the $\alpha$ chain; most of the correspondences between $\alpha$ and $\gamma$ chains are present also in the $\beta$ chain. The analysis of the correspondence can be extended to the $\delta$ chain, for which no ordering of the peptides has been provided, if one assumes that peptides with identical or very similar sequence or composition occupy a similar position along the $\beta$ and $\delta$ peptide chains. An upper limit for the extent of correspondence between the $\beta$ and the $\delta$ chains can be established in this way at 94%.

In 1957, when the structure of the human hemoglobins and the sequence of the hemoglobin chains were not yet elucidated, Itano considered the possibility that the genes for myoglobin—Hb-A, Hb-F, and Hb-$A_2$—were derived from a primitive gene through a process of gene duplication and independent evolution of the duplicated genes. This hypothesis has been examined in greater detail by Ingram (1961a) on the basis of the accumulating evidence of similarities in structure and amino acid sequence among the human hemoglobin chains.

Ingram (1961a) developed an evolutionary scheme which accounts for the similarities between the hemoglobin chains. In the course of evolution a gene or a chromosomal fragment, carrying a hemoglobin gene, has undergone duplication. Mutations of the two initially identical genes result in single or multiple amino acid substitutions, or possibly in inversions of the amino acid sequence of part of the peptide chains. These events are favored or discarded in the course of natural selection. The duplicated genes then evolve independently, under the selective pressure of the environment, and may be separated by chromosomal inversions or translocations. Several duplications followed by independent evolution would account for the presence of the four hemoglobin chains, for which an independent genetic control has been shown. The primary sequences of the $\alpha$, $\beta$, $\gamma$, and $\delta$ peptide chains have been shown to be determined by different structural genes; these genes undergo mutational events which affect only the corresponding peptide chains (see Section IV, A, E, and F). On the contrary, the peptide chains of the minor component Hb-$A_3$ have been shown to be not under independent genetic control. The chemical studies on Hb-$A_3$ (Muller, 1961) provide evidence that the peptide chains of Hb-A and of Hb-$A_3$ are under the control of the same structural genes and have an identical primary sequence.

### E. Hb-$A_3$

This minor component migrates in electrophoresis at pH 8.6 faster than Hb-A, but is not completely resolved from Hb-A in starch block electrophoresis (Kunkel and Wallenius, 1955). Considerable evidence was obtained by these authors that Hb-$A_3$ is a heterogeneous component, which may be further resolved into at least two fractions.

Hb-$A_3$ is most likely a derivative of Hb-A (Kunkel and Wallenius, 1955); Hb-$A_3$ increases in amount in old samples, particularly if Hb-A is present as methemoglobin. The rate of labeling of Hb-$A_3$ after $Fe^{59}$ administration was studied by Kunkel and Bearn (1957); Hb-$A_3$ showed a lower specific activity than Hb-A or Hb-$A_2$. This and several other observations have shown that the Hb-$A_3$ fraction, which can be further resolved by column chromatography (Schnek and Schroeder, 1961), contains a mixture of by-products of Hb-A, formed by alterations of the hemoglobin molecule during the aging of the red cells.

The fingerprint of Hb-$A_3$ isolated by column chromatography has shown a single difference from the fingerprint of Hb-A (Muller, 1961). A peptide reacting to the ninhydrin and sulfur test was present only in the fingerprints of Hb-$A_3$ and not in the fingerprints of purified Hb-A;

this peptide was shown to be glutathione on the basis of the amino acid composition and of chromatographic and electrophoretic mobility (Muller, 1961). Glutathione is not an integral part of the Hb-$A_3$ molecule, although it is rather stably attached; it is indeed released upon acid or heat denaturation but it is not removed by extensive dialysis (Muller, 1961). The type of chemical bond established between glutathione and hemoglobin is not known; aging of hemoglobin is required to bind glutathione, present in high concentration in red cells, to the hemoglobin molecule. The heterogeneity of the Hb-$A_3$ component can possibly be explained by the binding of a different number of glutathione molecules to the hemoglobin molecule. Glutathione has also been found in one of the minor components isolated from cord blood hemoglobin (Muller, 1961). Undoubtedly, many instances of heterogeneity of hemoglobins are likely to be explained by similar secondary modifications of the hemoglobin molecule. Environmental factors probably play a decisive role in modifying the hemoglobin molecule.

### F. EMBRYONIC HEMOGLOBIN

A hemoglobin different from Hb-F has been reported in very early fetuses. This hemoglobin has been designated primitive or embryonic type hemoglobin (Allison, 1955). Drescher and Künzer (1954) have reported a hemoglobin with intermediate alkali resistance in human fetuses 7–12 weeks of age. Halbrecht and Klibansky (1956) have reported a hemoglobin component with lower electrophoretic mobility than Hb-F, but with an Hb-F type UV spectrum in very early fetuses. Huehns *et al.* (1961a) have reported two hemoglobins with different electrophoretic mobilities in a survey of human fetus hemoglobins; these components were observed in the smallest fetuses examined, which had a crown-rump measurement of only 3.5 cm. No chemical characterization of these hemoglobins has, however, been reported.

The evidence for the existence of an embryonic hemoglobin is still unsatisfactory, since it has not been supported by careful chemical investigations. These investigations may rule out the possibility that the hemoglobins observed in very early fetuses correspond to artifacts or that they may in some instances represent genetic variants of Hb-F. It has been suggested by Matsuda *et al.* (1960), who failed to observe any hemoglobin different from Hb-F in human fetuses, that the "embryonic" hemoglobin is an artifact of the deterioration of Hb-F.

## IV. ABNORMAL HEMOGLOBINS

Abnormal hemoglobins are defined as the hemoglobins which differ in their physical and chemical properties from the normal components.

They are, with few exceptions, which will be considered separately, the direct expression of mutations which are inherited as autosomal genes.

The first hemoglobin abnormality reported in the literature was associated with congenital methemoglobinemia (Hörlein and Weber, 1948). These authors found that the hemoglobin from patients affected by this disease showed an abnormal absorption spectrum when in the form of acid methemoglobin; the disease was transmitted as a dominant factor through four generations. The first abnormal hemoglobin electrophoretically different from Hb-A was found by Pauling *et al.* (1949) in sickle-cell anemic patients. Hereditary transmission of this disease has been recognized for a long time, but until Neel's investigation in 1949 it was not clearly understood that the sickling gene is responsible for the mild abnormality known as sickle-cell trait in heterozygotes and for the severe sickle-cell anemia in homozygotes. Pauling *et al.* (1949) reached similar conclusions independently and were able to show that no Hb-A is present in homozygotes and that the heterozygotes have both Hb-A and the abnormal hemoglobin (Hb-S) in the approximate ratio 60:40.

These fundamental observations stimulated several investigators to examine patients affected by other forms of hereditary anemia and normal individuals, as well, for the presence of abnormal hemoglobins. Since 1950 many other abnormal hemoglobins have been discovered: Hb-C by Itano and Neel (1950), Hb-D by Itano (1951), Hb-E by Itano *et al.* (1954), and several others more recently (see Lehmann, 1960; Rucknagel and Neel, 1961). The enumeration of all the abnormal hemoglobins reported in recent years is outside the purpose of this review. Only short descriptions of few abnormal hemoglobins, which are of particular genetic or chemical interest, will be given in this review. Techniques involved in the detection of the abnormal hemoglobins have been reviewed by Lehmann (1960) and by Rucknagel and Neel (1961). The problem of the nomenclature of abnormal hemoglobins has become increasingly difficult. The letters of the alphabet were assigned in the order of their discovery to the various types of abnormal hemoglobins, with the exceptions of the letters F and S, reserved to fetal and sickle-cell hemoglobin. Since the abnormal hemoglobins discovered have outnumbered the letters of the alphabet, it has become necessary to name the newly discovered abnormal hemoglobins after the place where they have been found, following the example of Robinson *et al.* (1956). The first abnormal hemoglobins discovered were variants of the normal adult hemoglobin, Hb-A. Subsequently abnormal hemoglobins have been observed that have been shown to be variants of Hb-F and of Hb-$A_2$. These abnormal hemoglobins will be considered separately from the abnormal forms of Hb-A.

## A. Hb-A VARIANTS—CHEMICAL STUDIES

The abnormal forms of Hb-A represent the largest and best known group of abnormal hemoglobins. Most of the chemical studies have been carried out on variants of Hb-A and our knowledge of the inheritance of the hemoglobin genes and of the population distribution of the

TABLE II

Chemical Studies of the Hb-A Variants

| Charge difference[a] | Hemoglobin | Abnormal chain and tryptic peptides[b] | Amino acid substitution | Position | Reference |
|---|---|---|---|---|---|
| −4 | I | $\alpha$Tp III | Lys → Asp | 16 | Murayama and Ingram (1959) |
| −2 | K | $\alpha$ | | | Gammack *et al.* (1961) |
| | J | $\beta$ | | | Itano and Robinson (1959) |
| | $J_{Trinidad}$<br>$J_{Ireland}$<br>$J_{Jamaica}$ | $\beta$ | | | Gammack *et al.* (1961) |
| | $J_{Baltimore}$ | $\beta$Tp II | | | C. Baglioni, unpublished |
| | Norfolk | $\alpha$Tp VII | Gly → Asp | 57 | Baglioni (1962a) |
| | N | $\beta$ | | | Gammack *et al.* (1961) |
| | R (= Durham-1) | $\beta$Tp I | | | Chernoff and Liu (1961) |
| 0[c] | $M_{Boston}$ | $\alpha$Tp VII | His → Tyr | 58 | Gerald and Efron (1961) |
| | $M_{Saskatoon}$ | $\beta$Tp VII | His → Tyr | 63 | Gerald and Efron (1961) |
| | $M_{Milwaukee-1}$ | $\beta$Tp IX | Val → Glu | 67 | Gerald and Efron (1961) |
| | $M_{Milwaukee-2}$ | $\beta$ | | | Gerald and Efron (1961) |
| | $M_{Iwate}$ | $\alpha$Tp IX | | | Gerald and Efron (1961) |

[a] The charge difference with Hb-A at pH 8.6 is the theoretical charge difference based on the amino acid substitution and expressed in arbitrary units.

[b] For the peptide nomenclature see Table I and Gerald and Ingram (1961).

[c] The Hb-M's are observed as methemoglobins ($Fe^{+++}$); the formation of internal complexes between the iron and the substituting amino acids neutralizes the charge difference with oxy-Hb-A ($Fe^{++}$). The Hb-M's show a slightly different electrophoretic mobility compared to Hb-A (see Gerald and Efron, 1961).

*Table continued*

TABLE II* (*Continued*)

| Charge difference[a] | Hemoglobin | Abnormal chain and tryptic peptides[b] | Amino acid substitution | Position | Reference |
|---|---|---|---|---|---|
| +2 | $D_{\text{Punjab}}$ (= $D_{\text{Cyprus}}$) | $\beta$Tp XIII | Glu → Glu-$NH_2$ | 121 | Baglioni (1962b) |
| | $D_\alpha$ | $\alpha$Tp IV | | | Benzer *et al.* (1958) |
| | $D_\beta$ | $\beta$Tp III | | | Benzer *et al.* (1958) |
| | $D_{\text{Frankfurt}}$ | $\beta$ | | | Gammack *et al.* (1961) |
| | D | $\beta$ | | | Chernoff and Liu (1961) |
| | D | $\beta$ | | | Itano and Robinson (1959) |
| | $G_{\text{San José}}$ | $\beta$Tp I | Glu → Gly | 7 | Hill and Schwartz (1959) |
| | $G_{\text{Philadelphia}}$ (= $G_{\text{Azakuoli}}$) | $\alpha$Tp IX | Asp-$NH_2$ → Lys | 68 | Baglioni and Ingram (1961) and C. Baglioni, unpublished |
| | $G_{\text{Honolulu}}$ | $\alpha$Tp IV | Glu → Glu-$NH_2$ | 30 | Swenson *et al.* (1962) |
| | $G_{\text{Ibadan}}$<br>$G_{\text{Bristol}}$ | $\alpha$ | | | Gammack *et al.* (1961) |
| | $G_{\text{Chinese}}$ | $\beta$ | | | Gammack *et al.* (1961) |
| | L | $\beta$ | | | Gammack *et al.* (1961) |
| | P | $\beta$ | | | Gammack *et al.* (1961) |
| | Q | $\alpha$ | | | Gammack *et al.* (1961) |
| | S | $\beta$Tp I | Glu → Val | 6 | Hunt and Ingram (1959) |
| | Stanleyville I | $\alpha$Tp IV | | | Chernoff and Liu (1961) |
| | Zürich | $\beta$Tp VII | His → Arg | 63 | Muller and Kingma (1961) |

* For footnote to table see p. 423.

abnormal hemoglobin genes is based on the study of the Hb-A variants. The Hb-A variants can be roughly subdivided in four groups according to their electrophoretic mobility; this classification has been found useful in classifying new types of abnormal hemoglobins. In Table II are indicated the abnormal hemoglobin types discovered to date that have been chemically studied and shown to be variants of Hb-A.

TABLE II (*Continued*)

| Charge difference[a] | Hemoglobin | Abnormal chain and tryptic peptides[b] | Amino acid substitution | Position | Reference |
|---|---|---|---|---|---|
| +4 | C | βTp I | Glu → Lys | 6 | Hunt and Ingram (1959) |
| | E | βTp III | Glu → Lys | 26 | Hunt and Ingram (1959) |
| | $O_{Arabia}$ | βTp XIII | Glu → Lys | 121 | Baglioni and Lehmann (1962) |
| | $O_{Indonesia}$ | αTp XII | Glu → Lys | 116 | Baglioni and Lehmann (1962) and C. Baglioni, unpublished |

The chemical study of the abnormal hemoglobins by means of the fingerprinting method of analysis was initiated by Ingram (1956). The fingerprints of Hb-A and Hb-S showed that only one peptide in the Hb-S fingerprint occupies a position different from that in Hb-A fingerprints (see Fig. 2). After tryptic hydrolysis of the human hemoglobins a few large tryptic peptides, accounting for approximately ¼ to ⅓ of the molecule, are insoluble at neutral pH. These peptides are removed by centrifugation from the tryptic digest of the human hemoglobins prior to fingerprinting (Ingram, 1958). That portion of the hemoglobins which contains the peptides insoluble at neutral pH has been designated the "core" of the molecule (Ingram, 1958). The "core" can be examined by fingerprinting after chymotryptic digestion (Hunt and Ingram, 1958a). Hunt and Ingram (1958a) have shown that the fingerprints of the chymotryptic digest of the Hb-A and Hb-S "cores" are identical.

The isolation and the analysis of the peptide altered in Hb-S have allowed Ingram (1958) to establish that Hb-S differs from Hb-A by only one amino acid: a *valine* in Hb-S is substituted for a *glutamic acid* residue in Hb-A. This fundamental observation which revealed the alteration caused by a mutant gene at the molecular level undoubtedly was a landmark in the history of biology.

The next abnormal hemoglobin analyzed was Hb-C (Hunt and Ingram, 1958b); the same peptide displaced in Hb-S fingerprint is missing in Hb-C fingerprint, while two new peptides appear. The chemical analysis of these peptides (Hunt and Ingram, 1960) showed that the same residue substituted by a valine in Hb-S, namely, a glutamic acid, is replaced in Hb-C by a lysine residue. This new lysine offers a point of attack to trypsin and two peptides appear instead of the altered one in Hb-C fingerprints.

The known allelism of the Hb-S and Hb-C genes (Ranney, 1954) was thus confirmed at a strictly molecular level. These results established a correlation between the abnormal hemoglobins genes and the linear structure of the peptide chains, the synthesis of which is controlled by these genes.

Several other abnormal hemoglobins were analyzed in a similar way by Ingram and by others (see Table II). From the chemical studies of the peptide chains of Hb-A (Ingram, 1959a), it is known which peptides in the fingerprinting of this hemoglobin belong to the $\alpha$ chain and which belong to the $\beta$ chain (see Table I). By fingerprinting the abnormal hemoglobins it was possible to determine the peptide chain altered in these hemoglobins; i.e., peptides altered in fingerprints of Hb-S, Hb-E, and Hb-$\mathrm{D}_{\mathrm{Punjab}}$ were found to belong to the $\beta$ chain, peptides altered in Hb-I, Hb-Norfolk, and Hb-$\mathrm{G}_{\mathrm{Philadelphia}}$ were found to belong to the $\alpha$ chain. The alterations of the Hb-A variants have been assigned to one or the other of the Hb-A peptide chains by this method of analysis or by recombination experiments (Itano and Robinson, 1959; see Section IV,D). The chemical formula of an abnormal hemoglobin can be written by indicating with the hemoglobin designation the peptide chain that is altered, i.e., Hb-S = $\alpha_2{}^{\mathrm{A}}\,\beta_2{}^{\mathrm{S}}$, Hb-C = $\alpha_2{}^{\mathrm{A}}\,\beta_2{}^{\mathrm{C}}$, Hb-I = $\alpha_2{}^{\mathrm{I}}\,\beta_2{}^{\mathrm{A}}$.

Ingram (1959a) suggested a one gene-one peptide chain relationship for the Hb-A peptide chains. In this view each gene controls the synthesis of one type of peptide chain. The structural genes determining the amino acid sequence of the peptide chains are designated with the same symbols used to indicate the corresponding peptide chains; namely the gene $\alpha^{\mathrm{A}}$ carries the information for the synthesis of $\alpha^{\mathrm{A}}$ chains and the gene $\beta^{\mathrm{A}}$ for the synthesis of $\beta^{\mathrm{A}}$ chains. We can subdivide the Hb-A variants into two categories: the $\alpha$ chain and the $\beta$ chain variants. The Hb-S gene is defined as $\beta^{\mathrm{S}}$, the Hb-I gene as $\alpha^{\mathrm{I}}$, and so on.

One of the most significant results of the chemical analysis of abnormal hemoglobins is the discovery by Benzer *et al.* (1958) of the heterogeneity of hemoglobins, which are indistinguishable by means of physical methods of analysis. Three samples of Hb-D obtained from individuals living in different countries were analyzed and found to carry different peptide changes. These three Hb-D's are thus distinct proteins and, more important, they are certainly produced by three different mutant genes. The importance of such observation should be fully evaluated: the validity of genetic, ethnological, or anthropological correlations established in the absence of a biochemical analysis at the ultimate molecular level is in fact shown to be questionable. The heterogeneity of hemoglobins having identical electrophoretic and chromatographic characteristics has also been demonstrated for Hb-G

(Hill *et al.*, 1960; Baglioni and Ingram, 1961; Swenson *et al.*, 1962) and for Hb-0 (Baglioni and Lehmann, 1962). Three types of Hb-G carrying different peptide alterations have been described and possibly several other types may exist among the several Hb-G's reported in the literature. Two types of Hb-0 characterized by different alterations have been described (Baglioni and Lehmann, 1962). The heterogeneity of hemoglobin variants, which are indistinguishable on the basis of their physicochemical behavior, seems to be a rather general phenomenon. It is quite likely that several other examples of heterogeneity will be found in the near future.

All the abnormal hemoglobins examined have shown a single amino acid difference with Hb-A. Any single amino acid substitution can only produce one of four different types of electrophoretically altered proteins. If we arbitrarily assign the charge 0 to the normal protein, the substitution of a neutral amino acid by an acidic or a basic amino acid, as well as the substitution of a basic or acidic amino acid by a neutral amino acid, will result in a charge alteration of $-1$ or $+1$, respectively. The substitution of a basic amino acid by an acidic or of an acidic by a basic will result in charge alteration of $-2$ or $+2$, respectively. The charge difference is expressed in arbitrary units and should be multiplied by 2 when referred to hemoglobin, since two altered peptide chains are present in the molecule.

We know many more than four electrophoretically different abnormal hemoglobins, but the abnormal hemoglobins can be roughly grouped in five classes on the basis of their electrophoretic mobility (see Table II). A small electrophoretic difference exists at pH 8.6 between Hb-$O_{\text{Arabia}}$, Hb-C, and Hb-E, in which a lysine substitutes for a glutamic acid. These hemoglobins should have the same electrophoretic charge if the charge difference with Hb-A results only from the type of amino acid substitution. The small charge differences between these hemoglobins are probably due to different contributions of the residues to the ionic charge of the protein. The charge contribution is presumably dependent on the place occupied by a residue in the spatial configuration of the protein. The charge of ionizable groups may be modified by interaction with neighboring residues.

There seems to be no correlation between the type of amino acid involved in substitutions, apart from its charge, and the change in electrophoretic mobility. The Hb-G's, for instance, present different substitutions: lysine for glutamic in Hb-$G_{\text{San José}}$ (Hill *et al.*, 1960), lysine for asparagine in Hb-$G_{\text{Philadelphia}}$ (Baglioni and Ingram, 1961), and glutamine for glutamic in Hb-$G_{\text{Honolulu}}$ (Swenson *et al.*, 1962).

The electrophoretic behavior of the abnormal hemoglobins is discussed here only with reference to paper or starch-gel electrophoresis at pH 8.6, which are most commonly used. It should be pointed out that the abnormal hemoglobins may show a charge difference with Hb-A variable with the pH and the nature of the electrophoretic medium. The behavior of some abnormal hemoglobins in agar-gel electrophoresis is completely different from their behavior in starch-gel electrophoresis at similar pH's (see review by Gratzer and Beaven, 1961); i.e., Hb-$D_{Punjab}$ and Hb-E are not resolved from Hb-A by agar-gel electrophoresis at pH 6.2, while they are resolved by other types of electrophoresis and by chromatography on ion exchangers. It seems rather difficult to explain the electrophoretic and chromatographic behavior of the abnormal hemoglobins in terms of charge difference with Hb-A only; interactions at the level of the tertiary structure of the hemoglobin molecule may possibly account for some of the discrepancies observed. The charge of the amino acids involved in substitutions and their position in the tridimensional configuration of the protein determine the electrophoretic and chromatographic behavior of the abnormal hemoglobins.

### 1. *The Distribution of Amino Acid Substitutions in Hb-A Variants*

By determining the composition and the sequence of the peptides altered in abnormal hemoglobins, it has been possible to locate the residues substituted in the Hb-A variants along the sequences of the Hb-A peptide chains. In this way one can draw a map of the chemical changes caused by known mutations of the hemoglobin genes (Fig. 5). This is not a genetic map, but since the amino acid sequence of the peptide chains is probably parallel or colinear with the nucleotide sequence of the corresponding genes, it may correspond to a genetic map of the mutations in the hemoglobin gene. Intragenic recombination is a rare event and there is no doubt that the chances of observing such an event in humans are ridiculously small; thus, there is no hope to prove the colinearity between the amino acid sequence of a hemoglobin chain and the genetic fine structure of the corresponding gene by means of recombination of alleles. Demonstration of colinearity between genetic map and amino acid sequence is now being sought in more favorable organisms, as in *Escherichia coli*, for the gene controlling the enzyme trytophan synthetase (Helinski and Yanofsky, 1962) and phosphatase (Garen *et al.*, 1961) or in bacteriophage for the gene controlling lysozyme (Dreyer, 1960).

In Fig. 5 is shown schematically the position along the peptide chains of Hb-A of the amino acids substituted in the abnormal hemoglobins. There is no particular clustering of substitutions in any region of the peptide chains. It is remarkable, however, that some residues

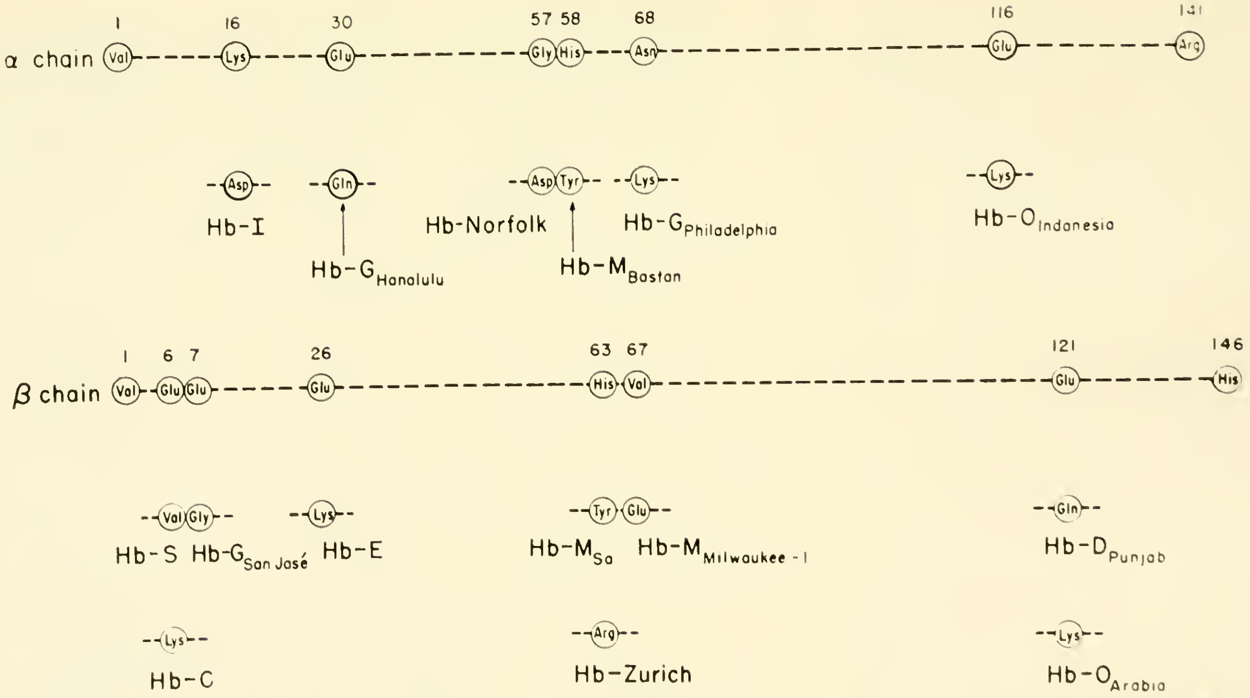

α chain
1
16
30
57 58
68
116
141
Val
Lys
Glu
Gly
His
Asn
Glu
Arg
Asp
Hb-I
Gln
Hb-G$_{Honolulu}$
Asp
Tyr
Hb-Norfolk
Hb-M$_{Boston}$
Lys
Hb-G$_{Philadelphia}$
Lys
Hb-O$_{Indonesia}$
β chain
1
6
7
26
63
67
121
146
Val
Glu
Glu
Glu
His
Val
Glu
His
Val
Gly
Hb-S
Hb-G$_{San José}$
Lys
Hb-E
Tyr
Glu
Hb-M$_{Sa}$
Hb-M$_{Milwaukee-1}$
Gln
Hb-D$_{Punjab}$
Lys
Hb-C
Arg
Hb-Zurich
Lys
Hb-O$_{Arabia}$

FIG. 5. Map of the amino acid substitutions in abnormal hemoglobins. The α and β peptide chains of Hb-A are schematically represented by dashed lines; the terminal amino acids and the ones substituted in abnormal hemoglobins are indicated inside small circles. The numbers above the circles indicate the position along the peptide chains of the corresponding amino acids. The substitutions found in the indicated abnormal hemoglobins are illustrated in correspondence with the amino acids of the normal peptide chains. Hb-$M_{Saskatoon}$ is indicated as Hb-$M_{Sa}$. For references see Table II.

appear to be involved more than once by an amino acid replacement: this is observed for residues 6, 63, and 121 of the β chain (see Table II for references). This does not necessarily imply that mutations occur more frequently in given regions of the hemoglobin genes (*hot spots*). Mutations probably occur in any position along the genes, with the consequent amino acid substitutions at various positions. However, abnormal hemoglobins carrying amino acid substitutions are probably synthesized at a normal rate only if the altered molecules are able to fold in the proper way. Therefore, we would not expect to observe substitutions involving certain amino acids because the corresponding mutations are particularly rare, but more likely because these mutations result in no production of hemoglobin. Higher levels of organization, the secondary and tertiary structures, which are dependent upon the primary structure, are in fact necessary for a protein to be stabilized.

### 2. *The Consequences of Amino Acid Substitutions on the Secondary and Tertiary Structures of Hemoglobin*

The secondary structure refers to the geometrical arrangement of the amino acids of a peptide chain in helical and non-helical regions; the tertiary structure refers to the final folded configuration of a peptide chain. We know very little about the alterations of the secondary and tertiary structures in abnormal hemoglobins. Watson and Kendrew (1961) have recently suggested that most of the residues involved in the amino acid substitutions of the abnormal hemoglobins stick outward from the three-dimensional configuration of the hemoglobin chains, so that these substitutions would not be expected to affect the tertiary structure considerably.

Watson and Kendrew (1961) have carefully compared the amino acid sequence of the sperm whale myoglobin with that of the human hemoglobin chains. The tertiary structure of the sperm whale myoglobin has been resolved by X-ray crystallography to a very high degree of precision by Kendrew and his collaborators (Kendrew *et al.*, 1961); of the 1200 atoms contained in the myoglobin molecule, excluding the hydrogens, some 75% have been localized without ambiguity in a three-dimensional model of this protein. Perutz *et al.* (1960) have shown by X-ray crystallography that the tertiary structure of the α and β chains of horse hemoglobin closely resembles that of the sperm whale myoglobin. Watson and Kendrew (1961) also have assumed that the secondary and tertiary structures of the human hemoglobin chains closely resemble that of the sperm whale myoglobin and have ideally aligned the hemoglobin peptide chains along the model of the sperm whale myoglobin. By this comparison most of the residues involved in substitutions in the abnormal hemoglobins have been found to lie on the surface of the tridimensional structure of the hemoglobin peptide chains. Few notable exceptions to this rule are represented by the amino acid residues involved in substitutions in the Hb-M's (Gerald and Efron, 1961).

The characteristic abnormality of this group of Hb-A variants is an unusual difficulty in reducing the ferric form of the iron atom in the heme group to the ferrous form. The Hb-M genes are responsible for the dominant hereditary form of methemoglobinemia; the recessive hereditary form of this disease is due to the absence or alteration of the enzyme diphosphopyridine nucleotide diaphorase (Scott and Griffith, 1959). The presence of Hb-M is associated with chronic cyanosis, since this hemoglobin is normally in the oxidized form as methemoglobin. Five samples of Hb-M from unrelated individuals have been examined by Gerald and Efron (1961). In three of these hemoglobins the amino acid substitution has been detected and localized along the hemoglobin chains.

Tyrosine has been found to substitute for histidine in homologous regions of the $\alpha$ chain in Hb-$M_{\text{Boston}}$ (Fig. 6), and of the $\beta$ chain in Hb-$M_{\text{Saskatoon}}$ (Fig. 7). On the basis of the similarities with the tertiary structure of the sperm whale myoglobin, Watson and Kendrew (1961) have suggested that these regions of the human hemoglobin chains are situated immediately opposite the sixth coordination position of the iron atom in the heme. The heme is known to be located in the inner part of each folded chain, between the folds of the chain; it is most likely attached to a portion of the peptide chains opposite the one altered in Hb-$M_{\text{Boston}}$ and Hb-$M_{\text{Saskatoon}}$. If, because of a gene mutation, an amino acid with a reactive side chain is substituted for an amino acid with no such side chain in the vicinity of the heme group, then the reactive side chain may complex with the oxidized heme group to give a stable complex. This is an internal complex in which the reactive group of the amino acid acts as a ligand. The complex is quite stable and for this reason the Hb-M remains in the ferric form as methemoglobin (Gerald and Efron, 1961).

That part of the hemoglobin chains where the amino acid substitutions leading to Hb-M have been found is presumably in helical configuration (Watson and Kendrew, 1961). The pitch of the helix is about 100° per residue, so that the side chains of two adjacent amino acids point toward different directions at a 100° angle. Only residues four positions apart have their side chains pointing toward the same side at a 40° angle.

Other amino acid substitutions have been observed in the same region of the peptide chains where the Hb-$M_{\text{Boston}}$ and Hb-$M_{\text{Saskatoon}}$ substitu-

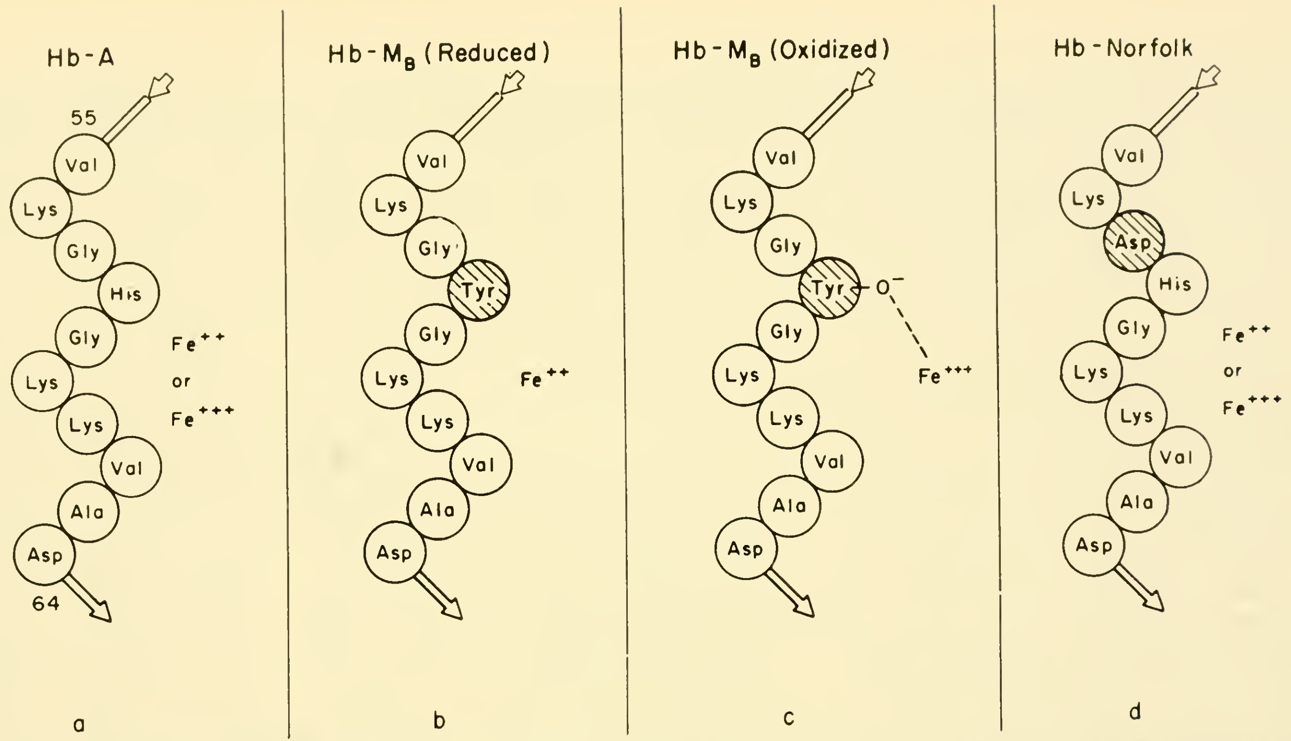

Hb-A
55
Val
Lys
Gly
His
Gly
Lys
Lys
Val
Ala
Asp
64
Fe⁺⁺
or
Fe⁺⁺⁺
a
Hb-M$_B$ (Reduced)
Val
Lys
Gly
Tyr
Gly
Lys
Lys
Val
Ala
Asp
Fe⁺⁺
b
Hb-M$_B$ (Oxidized)
Val
Lys
Gly
Tyr
O⁻
Gly
Lys
Lys
Val
Ala
Asp
Fe⁺⁺⁺
c
Hb-Norfolk
Val
Lys
Asp
His
Gly
Lys
Lys
Val
Ala
Asp
Fe⁺⁺
or
Fe⁺⁺⁺
d

FIG. 6. Schematic representation of the section of the *α* peptide chain from residue 55 to residue 64. (a) Hb-A. (b) Hb-M$_{\text{Boston}}$ in the reduced form. (c) Hb-M$_{\text{Boston}}$ in the oxidized form. (d) Hb-Norfolk. This section of the *α* peptide chain is thought to be in helical configuration and to be situated opposite to the heme group in the folded chain (Watson and Kendrew, 1961). Only the iron atom of the heme group is indicated in its reduced ($Fe^{++}$) or oxidized ($Fe^{+++}$) form. The helical configuration of this region of the peptide chain is schematically illustrated by the arrangement of the circles enclosing the amino acids. The side chains of amino acids four positions apart point approximately toward the same direction (see Section IV,E,2). Circles shaded in (b), (c), and (d) indicate the amino acid substitutions in the corresponding abnormal hemoglobins. The dashed line between Tyr and $Fe^{+++}$ in (c) illustrates the formation of a stable complex between the side chain of the tyrosine residue and the heme group (Gerald and Efron, 1961). For references to the amino acid substitutions in abnormal hemoglobins see Table II.

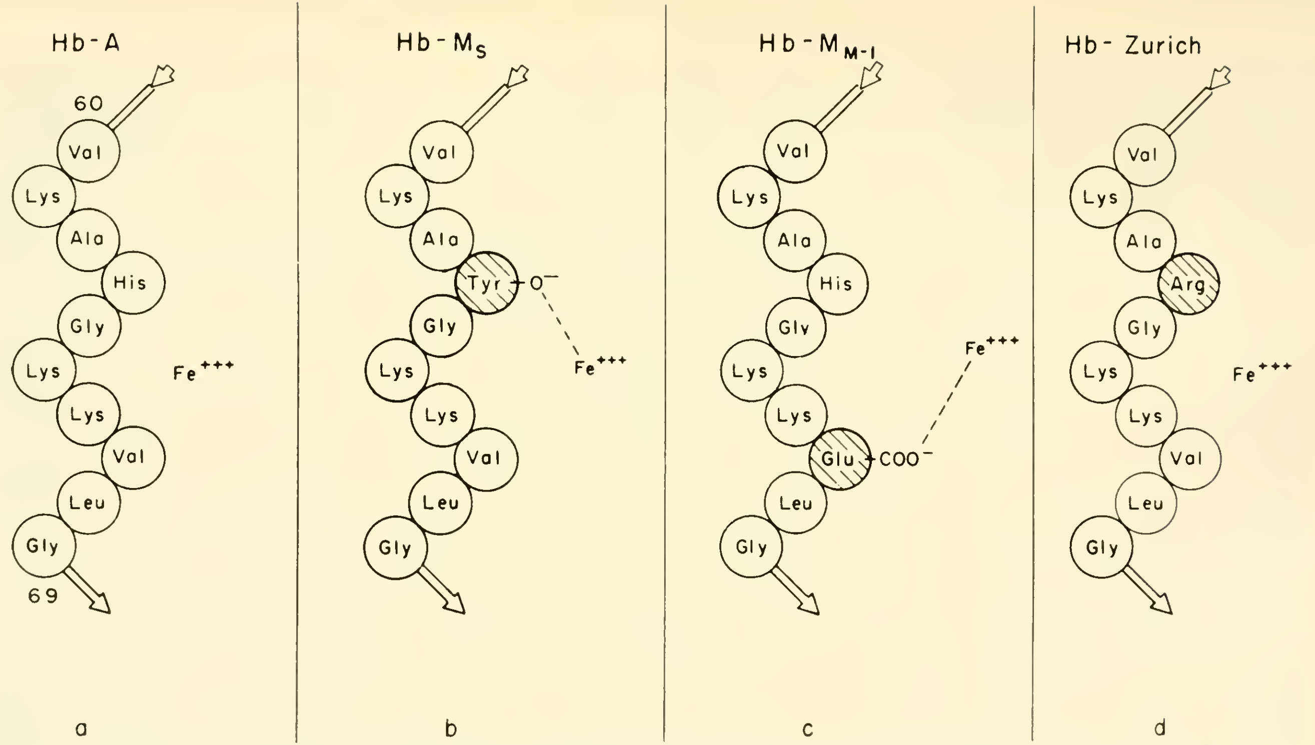

Hb-A
60
Val
Lys
Ala
His
Gly
Lys
Lys
Val
Leu
Gly
69
Fe+++
a
Hb-Ms
Val
Lys
Ala
Tyr
O−
Gly
Lys
Lys
Val
Leu
Gly
Fe+++
b
Hb-MM-I
Val
Lys
Ala
His
Gly
Lys
Lys
Glu
COO−
Leu
Gly
Fe+++
c
Hb-Zurich
Val
Lys
Ala
Arg
Gly
Lys
Lys
Val
Leu
Gly
Fe+++
d

FIG. 7. Schematic representation of the section of the $\beta$ peptide chain from residue 60 to 69. (a) Hb-A. (b) Hb-$M_{\text{Saskatoon}}$. (c) Hb-$M_{\text{Milwaukee-1}}$. (d) Hb-Zurich. This part of the $\beta$ peptide chain is thought to be in helical configuration and to be situated opposite to the heme group in the folded chain (Watson and Kendrew, 1961). Only the iron atom of the heme group is indicated in its oxidized form ($Fe^{+++}$). The helical configuration of this region of the peptide chain is schematically illustrated by the arrangement of the circles enclosing the amino acids. The side chains of amino acids four positions apart point approximately toward the same direction (see Section IV,E,2). Circles shaded in (b), (c), and (d) indicate the amino acid substitutions in the corresponding abnormal hemoglobins. The dashed lines between Tyr in position 63 in (b) or Glu in position 67 in (c) and $Fe^{+++}$ illustrate the formation of stable complexes between the side chains of the tyrosine or glutamic acid residues and the heme group (Gerald and Efron, 1961). For references to the amino acid substitutions in abnormal hemoglobins see Table II.

tions have been located, but none of these substitutions causes the formation of a stable methemoglobin by forming an internal complex as the Hb-M's do. This is the situation also when the substitution involves an amino acid with a reactive side chain, like Hb-Norfolk, where an aspartic acid substitutes for the glycine next to the histidine in Hb-$M_{Boston}$ (Baglioni, 1962a). The reactive side chain of the aspartic acid very likely does not point toward the heme group and cannot complex with the iron atom (see Fig. 6).

In the $\beta$ chain a different Hb-M ($M_{Milwaukee-1}$) has been found to have a glutamic acid substituting for a valine in position 67, four residues apart from the histidine in position 63, substituted in Hb-$M_{Saskatoon}$. The side chain of the glutamic acid probably points toward the heme and complexes with the iron atom.

The abnormal hemoglobin Hb-Zurich has recently been investigated by Muller and Kingma (1961); in Hb-Zurich the same histidine substituted by tyrosine in Hb-$M_{Saskatoon}$ is substituted by arginine. In Hb-Zurich apparently there is no interaction between the arginine residue in position 63 and the heme group (see Fig. 7). In order to have an interaction the requirement is not only the appropriate geometrical position of a residue in the secondary and tertiary configurations of the hemoglobin chain, but also the presence of a reactive side chain. The Hb-Zurich carriers show a puzzling phenomenon: severe hemolytic crisis and inclusion bodies in the red cells after sulfanilamide administration (Hitzig *et al.*, 1960). The presence of Hb-Zurich appears thus to be associated with a drug-induced inclusion body anemia; this abnormal hemoglobin is presumably altered following sulfanilamide administration, and precipitates, forming inclusion bodies.

### 3. *The Nucleotide Base Code and the Amino Acid Substitutions in Abnormal Hemoglobins*

Nirenberg and Matthaei (1962) have recently discovered that polyuridylic acid directs the synthesis of polyphenylalanine in a cell-free protein-synthesizing system, dependent upon the addition of messenger RNA. This indicated that the nucleotide sequence coding for phenylalanine in messenger RNA is made up of uridylic acid only. The effect of randomly ordered polyribonucleotides upon the incorporation of other amino acids has been investigated by Lengyel *et al.* (1961), by Matthaei *et al.* (1962), and by Speyer *et al.* (1962), in order to determine the composition of RNA coding units. The work of the above-mentioned authors has elucidated the four letter nucleotide code, which specifies the primary sequence of amino acids in proteins. This code is presumably a comma-less, non-overlapping code, as suggested by the genetic experiments of Crick *et al.* (1961). These authors have obtained evidence that the code is a three-letter, or a multiple of three letters, code. In agreement, the minimum number of nucleotides in coding units of copolymers of three different ribonucleotides, which stimulate the incorporation of some amino acids, is obviously three.

Triplet code letters have been experimentally obtained for 19 amino acids (Matthaei *et al.*, 1962; Speyer *et al.*, 1962). An experimental value for the glutamine coding unit is still unavailable, but the composition of this coding unit has been predicted from amino acid replacement data in nitrous acid mutants of tobacco mosaic virus (Speyer *et al.*, 1962). Smith (1962a) has examined the relationships between the proposed triplet code and the amino acid substitutions in polypeptides or proteins of known structure, pointing out that all the known amino acid replacements in the abnormal human hemoglobins are consistent with a single

TABLE III

THE AMINO ACID CODE[a]

| Amino acid | Code |
|---|---|
| Ala | CUG[b] |
| Arg | UGC |
| Asp-$NH_2$ | UCA, (AUA) |
| Asp | UGA |
| Cys | (GUU), (GUG) |
| Glu | UAG |
| Glu-$NH_2$ | UCG |
| Gly | UGG |
| His | UAC |
| Ileu | UUA |
| Leu | UCU, AUU, (UGU) |
| Lys | UAA |
| Met | GUA |
| Phe | UUU |
| Pro | CCU |
| Ser | CUU, (UCG) |
| Thr | CUA, UUC |
| Try | (GGU) |
| Tyr | UAU |
| Val | UUG |

[a] The code is constructed according to Smith (1962b), assigning to glutamic acid the base sequence UAG. The triplets composition is from Matthaei *et al.* (1962) and from Speyer *et al.* (1962). The sequence of the triplets in parenthesis cannot presently be established. This is one of the six possible codes, based on the permutations of the sequence for glutamic acid arbitrarily assigned.

[b] Abbreviations: A, C, G, and U are used for adenylic, cytidylic, guanylic, and uridylic nucleotide residues in RNA, respectively.

base change in the proposed code. Smith (1962b) has shown that if an arbitrary sequence is assigned to the triplet coding for a given amino acid (i.e., the base sequence for glutamic acid is assumed to be UAG), the sequence of the triplets coding for 18 amino acids can be established utilizing the known amino acid substitutions in abnormal hemoglobins and in nitrous acid mutants of tobacco mosaic virus (see Table III). It is possible to construct in this way six sets of codes, based on the six permutations of the sequence arbitrarily assigned; one of these is likely to represent the correct code in messenger RNA. However, for some amino acids two or three alternative coding units have been reported (Matthaei *et al.*, 1962; Speyer *et al.*, 1962). The code appears thus to be degenerate, as suggested also by the elegant genetic experiments of Crick *et al.* (1961).

In view of the degeneracy of the code it is of some interest to examine in detail instances of multiple amino acid replacements of the same residue. These replacements are evidently produced by different changes of the same coding units. Three double replacements have been found in the $\beta$-peptide chain of Hb-A (for references and explanations see Tables II and III):

$$\begin{tikzcd}
 & \beta^{S}6 & & \mathrm{Val} & & \mathrm{UUG} \\
\beta^{A}6 \arrow[ur] \arrow[dr] & & \mathrm{Glu} \arrow[ur] \arrow[dr] & & \mathrm{UAG} \arrow[ur] \arrow[dr] & \\
 & \beta^{C}6 & & \mathrm{Lys} & & \mathrm{UAA}
\end{tikzcd}$$

$$\begin{tikzcd}
 & \beta^{M}{}_{\mathrm{Sa}}\ 63 & & \mathrm{Tyr} & & \mathrm{UAU} \\
\beta^{A}\ 63 \arrow[ur] \arrow[dr] & & \mathrm{His} \arrow[ur] \arrow[dr] & & \mathrm{UAC} \arrow[ur] \arrow[dr] & \\
 & \beta^{\mathrm{Zurich}}\ 63 & & \mathrm{Arg} & & \mathrm{UGC}
\end{tikzcd}$$

$$\begin{tikzcd}
 & \beta^{D\ \mathrm{Punjab}}\ 121 & & \mathrm{Glu\text{-}NH_2} & & \mathrm{UCG} \\
\beta^{A}121 \arrow[ur] \arrow[dr] & & \mathrm{Glu} \arrow[ur] \arrow[dr] & & \mathrm{UAG} \arrow[ur] \arrow[dr] & \\
 & \beta^{O\ \mathrm{Arabia}}\ 121 & & \mathrm{Lys} & & \mathrm{UAA}
\end{tikzcd}$$

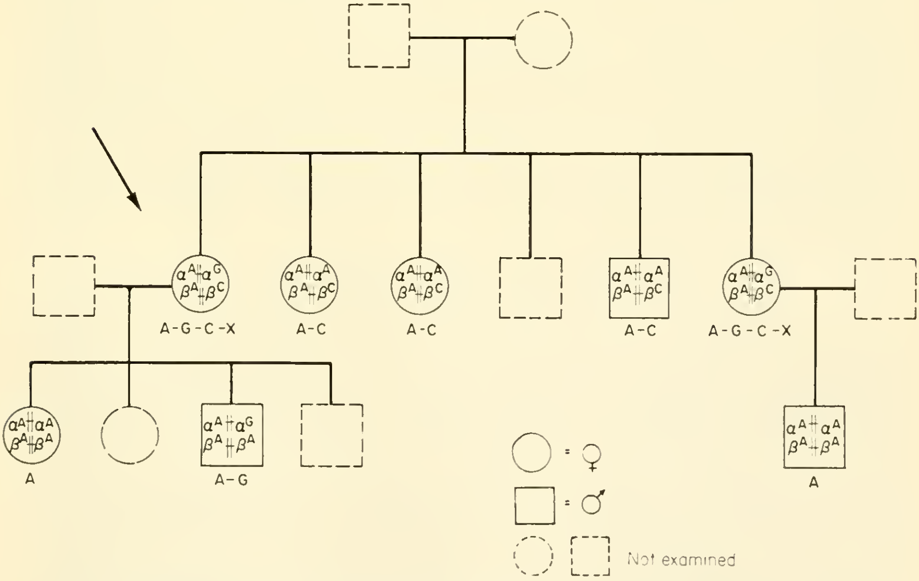
These replacements are in complete agreement with the proposed genetic code (Smith, 1962b).

### B. FORMAL GENETICS OF THE Hb-A VARIANTS

Mendelian inheritance of the abnormal hemoglobin genes has been shown for most of the Hb-A variants (Neel, 1956). The most pertinent information on the inheritance of the hemoglobin genes is obtained from marriages where one partner possesses two abnormal hemoglobins and the other none (Rucknagel and Neel, 1961). Ranney (1954) described marriages in which both Hb-C and Hb-S were observed in one partner; only children with either Hb-S or Hb-C were found among the offspring of these marriages. The allelism of the Hb-C and Hb-S genes was confirmed a few years later by the chemical investigation of Hunt and Ingram (1958b); the same amino acid residue was found to be substituted in these two abnormal hemoglobins.

One marriage supporting allelism of the Hb-E and Hb-S genes has been described by Aksoy and Lehmann (1957). Very few marriages involving a partner with two chemically characterized abnormal hemoglobins have been reported. No individuals having two hemoglobins abnormal in the α chain have thus far been reported.

The low frequency with which the hemoglobins abnormal in the α chain are observed in populations suggests very small probability that



$\alpha^A \alpha^G$ $\beta^A \beta^C$
A-G-C-X
$\alpha^A \alpha^A$ $\beta^A \beta^C$
A-C
$\alpha^A \alpha^A$ $\beta^A \beta^C$
A-C
$\alpha^A \alpha^A$ $\beta^A \beta^C$
A-C
$\alpha^A \alpha^G$ $\beta^A \beta^C$
A-G-C-X
$\alpha^A \alpha^A$ $\beta^A \beta^A$
A
$\alpha^A \alpha^G$ $\beta^A \beta^A$
A-G
$\alpha^A \alpha^A$ $\beta^A \beta^A$
A
= ♀
= ♂
Not examined

FIG. 8. Family tree of the propositus (indicated by the arrow) with four hemoglobins. (From Atwater *et al.*, 1960a). The genotype of the individuals examined is illustrated. The capital letters designate the hemoglobin types observed in the corresponding individuals. The propositus and one sister are doubly heterozygous.

critical marriages involving individuals with two of these hemoglobins can be encountered.

Allelism for most of the abnormal hemoglobin genes has been postulated because of the absence of Hb-A in individuals having two abnormal hemoglobins. Individuals having Hb-S and Hb-$D_{Punjab}$ or having Hb-S and Hb-N do not show any Hb-A, even in trace amounts. The first exception to the observation that persons with two abnormal hemoglobins do not have any Hb-A was reported by Smith and Torbert (1958). These authors described the occurrence in a family of Hb-S and of a second abnormal hemoglobin (Hb-Hopkins-2); in all the members of this family showing both abnormal hemoglobins, Hb-A also was found to be present. Smith and Torbert (1958) reported a three-generation pedigree showing evidence for independent segregation of the two abnormal hemoglobin genes.

Atwater *et al.* (1960a) have reported the presence in two sisters of four hemoglobins; in the three children of the two sisters, presumably married to normal individuals, either no abnormal hemoglobins or only one abnormal hemoglobin was observed (Fig. 8). Other individuals with three (Cabannes and Portier, 1959) and four hemoglobins (Raper *et al.*, 1960) have been reported. The existence of two hemoglobin genes is shown by the segregation of the abnormal hemoglobins in the families reported by Smith and Torbert (1958) and Atwater *et al.* (1960a). Atwater *et al.* (1960a) suggested possible schemes of inheritance of these genes; they considered it most likely that these genes are located on two different chromosomes. The possibility that the two genes are located far apart on the same chromosome should also be considered; the small number of individuals observed does not provide definite information on the position of the two genes.

### C. GENETIC AND BIOCHEMICAL RELATIONSHIPS

Of particular interest is the presence in two members of the family reported by Atwater *et al.* (1960a) of four electrophoretically distinct hemoglobins. Baglioni and Ingram (1961) have investigated the chemical alterations of these four hemoglobins, which have been called Hb-A, Hb-G, Hb-C, and Hb-X, according to the order of migration in electrophoresis. The Hb-C from this family was found to be identical to the Hb-C described by Hunt and Ingram (1960) and consequently altered in the $\beta$ chain. The Hb-G (designated Hb-$G_{Philadelphia}$) was found to be altered in the $\alpha$ chain, where a lysine substitutes for an asparagine residue. Hb-X was found to contain the abnormal peptides characteristic of both Hb-C and Hb-$G_{Philadelphia}$; Hb-X is thus abnormal in both its $\alpha$ and $\beta$ peptide chains. A similar situation has been described by Itano and Robinson (1960) for the individuals of the family reported by Smith and Torbert (1958) with three hemoglobins. Hemoglobin Hopkins-2 has been found to be altered in the $\alpha$ chain by dissociation and reassociation experiments (Singer and Itano, 1959; see Section IV); the alteration of the second abnormal hemoglobin, Hb-S, observed in the family is known to reside in the $\beta$ chain. Hb-S and Hb-Hopkins-2 have altered electrophoretic mobilities; Hb-S has a +2 charge and Hb-Hopkins-2 a −2 charge with respect to Hb-A. The doubly abnormal hemoglobin $\alpha_2^{\text{Hopkins-2}}\,\beta_2^{\text{S}}$ happens to have the same charge as Hb-A. Itano and Robinson (1960) have shown that this hemoglobin is present in the individuals with three hemoglobins of the family reported by Smith and Torbert (1958).

The occurrence of hemoglobin molecules abnormal in both peptide chains allows one to make certain deductions about the mode of assembly of the hemoglobin molecule. The persons with four hemoglobins are doubly heterozygous, having the genotype $\alpha^{\text{A}}/\alpha^{\text{G}}\,\beta^{\text{A}}/\beta^{\text{C}}$, as suggested by Atwater *et al.* (1960a) or $\alpha^{\text{A}}/\alpha^{\text{Hopkins-2}}\,\beta^{\text{A}}/\beta^{\text{S}}$ (Itano and Robinson, 1960). Apparently each gene causes the manufacture of a peptide chain characteristic of it, so that four types of peptide chains are produced in these persons. It has to be assumed that each gene controls the synthesis of a dimer, rather than that of single chains. Hemoglobin molecules having two different $\alpha$ or $\beta$ chains, such as $\alpha^{\text{A}}\,\alpha^{\text{G}}\,\beta_2^{\text{A}}$ or $\alpha_2^{\text{A}}\,\beta^{\text{A}}\,\beta^{\text{S}}$, have never been observed (Itano, 1956). The fact that each gene controls the synthesis of a dimer may be a consequence of the way in which the hemoglobin molecules are synthesized. A possible explanation is that the chains have to dimerize in order to be released in solution. In this view each microsomal particle has the information for only one type of peptide chain or possibly several copies of the same peptide chain; the newly synthesized chains are scarcely soluble and are released in solution only upon dimerization. The different types of dimers present assemble in a random fashion to complete hemoglobin molecules, leading to all the possible combinations (see Fig. 9).

An alternative explanation is that each chain associates with an identical partner because of steric hindrances, which prevent association of non-identical partners. It is rather difficult to hold such an explanation in view of the rather limited alterations in three-dimensional structure that certain Hb-A variants seem to have. Hb-A and Hb-$\text{D}_{\text{Punjab}}$, for instance, differ by only one amide group; a glutamic acid in Hb-A is changed to glutamine in Hb-$\text{D}_{\text{Punjab}}$ in position 121 of the $\beta$ chain (Baglioni, 1962b). It seems unlikely that this limited alteration could prevent association of the $\beta^{\text{A}}$ chains with the $\beta^{\text{D}_{\text{Punjab}}}$ chains.

According to the former explanation for the specificity of the hemoglobin chain dimerization, the existence of a hemoglobin molecule having two different types of $\alpha$ or $\beta$ chains should be possible. No such molecule

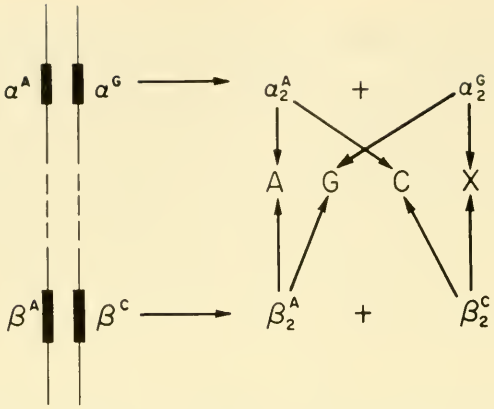


FIG. 9. Schematic explanation of the presence of four hemoglobins in doubly heterozygous individuals (Atwater *et al.*, 1960a; Baglioni and Ingram, 1961). The loci controlling the synthesis of the $\alpha$ and $\beta$ peptide chains are represented on the left. The arrows show the different dimers synthesized under the control of the $\alpha$ and $\beta$ alleles and the random combination of these dimers in the last stage of the assembly of the hemoglobin molecule. The letters A, G, C, and X indicate the four hemoglobin types synthesized. Hb-A results from the combination of $\alpha_2{}^A$ dimers with $\beta_2{}^A$ dimers, Hb-G from the combination of $\alpha_2{}^G$ dimers with $\beta_2{}^A$ dimers, Hb-C from the combination of $\alpha_2{}^A$ with $\beta_2{}^C$ dimers, and Hb-X from the combination of $\alpha_2{}^G$ with $\beta_2{}^C$ dimers.

has yet been reported to be produced by experiments of dissociation and recombination; however, the hemoglobin molecule is not dissociated to single chains in these experiments. (Hasserodt and Vinograd, 1959).

### D. DISSOCIATION AND RECOMBINATION OF HEMOGLOBIN

Field and O'Brien (1955) reported that hemoglobin dissociates reversibly at acid and alkaline pH's into subunits of molecular weight approximately half that of undissociated hemoglobin. Itano and Singer (1958) applied the acid dissociation method to the study of the dissociation and recombination of mixtures of hemoglobins A, S, and C. Singer and Itano (1959) and Vinograd *et al.* (1959) observed transfer of $C^{14}$-labeled subunits from Hb-A to Hb-C or Hb-S upon dissociation and recombination. When equivalent amounts of Hb-$A^{C^{14}}$ and Hb-C were recombined, one quarter of the radioactivity was found to be transferred to Hb-C. By these experiments Itano and his collaborators and Vinograd *et al.* (1959) have been able to show that different hemoglobins exchange subunits by a specific mechanism of dissociation and recombination, which results in the formation of hybrid molecules. However, formation of hybrid molecules containing two different $\alpha$ or $\beta$ chains does not occur. Recombination of hemoglobins altered in different chains provides a more direct and striking demonstration of the exchange of subunits upon recombination. Itano and Robinson (1959) recombined a mixture of Hb-I ($\alpha_2{}^I \beta_2{}^A$) and of Hb-S ($\alpha_2{}^A \beta_2{}^S$) and observed the formation of two new hemoglobins, one with the electrophoretic mobility of Hb-A ($\alpha_2{}^A \beta_2{}^A$) and the other, electrophoretically faster than Hb-A, resulting from the combination of $\alpha_2{}^I$ dimers with $\beta_2{}^S$ dimers. Observing the presence of four hemoglobins after recombination of two hemoglobins having different peptide chains, Itano and Robinson (1959) predicted that in individuals doubly heterozygous for two abnormal hemoglobins, a similar situation could be found. Their prediction was soon to be validated by the discovery of individuals with four hemoglobins (Atwater *et al.*, 1960a; Raper *et al.*, 1960).

Jones *et al.* (1959a) have used dissociation at alkaline pH and recombination, to determine the gross chemical structure of Hb-H (see Section IV,H,1). Recombination of Hb-H with Hb-$S^{C^{14}}$ ($\alpha_2{}^{A*} \beta_2{}^{S*}$) resulted in the formation of four components, which were identified as $\beta_4{}^A$, $\beta_4{}^{S*}$, $\alpha_2{}^{A*} \beta_2{}^A$, and $\alpha_2{}^{A*} \beta_2{}^{S*}$, thus showing that Hb-H is an abnormal hemoglobin made up of $\beta^A$ chains only.

Jones *et al.* (1959b) have shown that $\alpha_2{}^{A*}$ subunits can be transferred from Hb-$S^{C^{14}}$ ($\alpha_2{}^{A*} \beta_2{}^{S*}$) to Hb-F and that $\alpha_2$ subunits from Hb-F could associate with $\beta_2{}^A$ dimers produced by dissociation of Hb-H to give Hb-A ($\alpha_2{}^A \beta_2{}^A$); this indicates that the $\alpha_2$ dimers of Hb-A and Hb-F can be exchanged and are probably identical. Weatherall and Boyer (1962) have reported the formation of Hb-A and hemoglobin F/G ($\alpha_2{}^G \gamma_2{}^F$) upon acid recombination of Hb-F ($\alpha_2{}^A \gamma_2{}^F$) and of Hb-G ($\alpha_2{}^G \beta_2{}^A$).

The results of the recombination experiments have suggested (Itano and Robinson, 1959) a method for identifying the altered peptide chain of any electrophoretically abnormal hemoglobin. A hemoglobin in which the abnormal chain has been identified is recombined with an abnormal hemoglobin in which the abnormal chain is unknown. If the two hemoglobins are abnormal in different peptide chains, namely, if one has abnormal $\alpha$ chains and the other abnormal $\beta$ chains, new hemoglobin species, which are electrophoretically different will be produced upon recombination. If the two hemoglobins are abnormal in the same chain, no new hemoglobin species will be produced upon dissociation and recombination. A microscale method of analysis has been devised by Gammack *et al.* (1960) to determine by dissociation and recombination the altered peptide chain of an abnormal hemoglobin. By this method of analysis Gammack *et al.* (1961) were able to individuate the abnormal peptide chain of several abnormal hemoglobins. The dissociation and recombination analysis has also been applied to the study of other human hemoglobins and of animal hemoglobins.

Huehns and Shooter (1961) have shown by recombination that Hb-$A_2$ dissociated into $\alpha_2^A$ and $\delta_2^{A_2}$ subunits and that the $\alpha_2^A$ dimers can combine with $\beta_2^A$ dimers to give a hemoglobin electrophoretically identical to Hb-A. In order to obtain recombination of Hb-$A_2$, prolonged exposure to acid pH is necessary and equilibrium in the exchange of subunits is not reached after 72 hours. The rate of exchange of Hb-$A_2$ subunits is considerably increased by dissociation in acid buffers of higher ionic strength (Weatherall and Boyer, 1961). These findings suggest that the chemical bonds between the subunits of Hb-$A_2$ are probably different from the ones involved in stabilizing the quaternary structure of Hb-A.

The ways in which the exchange of subunits occurs upon dissociation and recombination merits further experimentation. A contradiction exists between the fairly rapid change in molecular weight of the hemoglobin molecule at acid pH and the relatively long time required to reach equilibrium in the exchange of subunits (J. R. Vinograd, personal communication). Itano *et al.* (1959) considered two possible mechanisms to explain the results of dissociation and reassociation experiments: (*1*) Hemoglobin dissociates symmetrically into half-molecules, but only identical half-molecules recombine to re-form the original components; (*2*) hemoglobin dissociates asymmetrically into pairs of identical chains. Mechanism (*2*) is in agreement with the recombination experiments, while mechanism (*1*) would explain the rapid reduction in molecular weight of hemoglobin in acid solutions, which is not followed by recombination of subunits. The factors which may prevent aggregation of non-identical half-molecules obtained by symmetric dissociation are not understood at all. We are still far from having a clear-cut explanation of the physicochemical phenomena involved in dissociation and reassociation of hemoglobin molecules.

### E. Hb-F variants

Electrophoretically abnormal forms of Hb-F have been reported by Fessas and Papaspyrou (1957), by Vella (1959), and by Fessas *et al.* (1961). These variant forms of Hb-F have not been chemically investigated, so that they are still rather undefined. They are thought to have altered $\gamma$ peptide chains; this would prove that the $\gamma$ peptide chain is under independent genetic control. The occurrence of mutant forms of a $\gamma$ chain gene has not been shown experimentally; the existence of this gene is predicted by the one gene-one peptide chain hypothesis (Ingram, 1959a).

#### 1. *The $\alpha$ Chains of Hb-F*

Minnich *et al.* (1962) have reported the occurrence in the cord blood of two Negro babies of a hemoglobin with an electrophoretic mobility somewhat slower than that of Hb-D; this hemoglobin was found to be alkali-resistant and to have the same N-terminal amino acids as Hb-F. One of the parents of each baby was heterozygous for $Hb\text{-}D_{St. Louis}$. Minnich *et al.* (1962) postulated that the abnormal fetal hemoglobin resulted from the combination of $\alpha^{D_{St. Louis}}$ chains with $\gamma^F$ chains. Similar cases have been observed by Weatherall and Boyer (1962). The cord blood hemoglobins of a Negro baby have been studied by Weatherall and Baglioni (1962); column chromatography on IRC-50, according to Allen *et al.* (1958), showed the presence of four hemoglobins, identified by fingerprinting as Hb-F, $Hb\text{-}F/G_{Philadelphia}$ Hb-A, and $Hb\text{-}G_{Philadelphia}$ in the order of elution from the column. The alteration of $Hb\text{-}G_{Philadelphia}$ is known to reside in the $\alpha$ peptide chain (Baglioni and Ingram, 1961).

These results show that the $\alpha$ chain of Hb-A and Hb-F are under the control of the same $\alpha$ gene. The $\alpha$ peptide chains of Hb-A and Hb-F have been shown to be chemically indistinguishable (Hunt, 1959; Schroeder *et al.*, 1961). Normal $\alpha_2{}^A$ dimers and abnormal $\alpha_2{}^G$ dimers are produced in

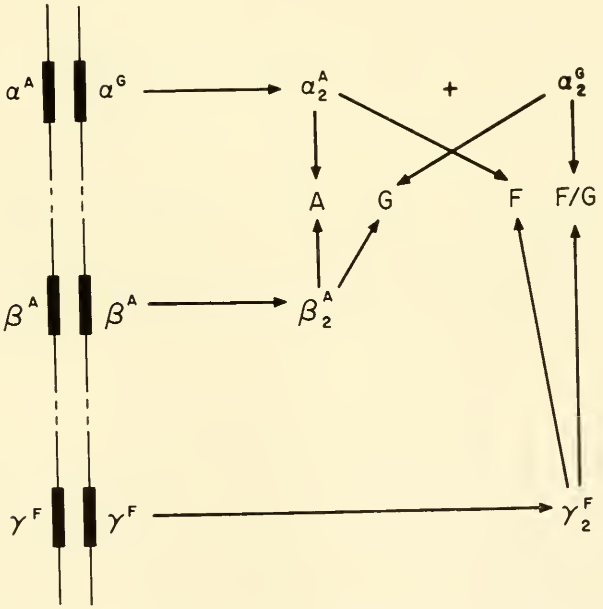


FIG. 10. Schematic explanation of the presence of four hemoglobins in fetuses or newborns heterozygous for an $\alpha$ chain mutation (Minnich *et al.*, 1962; Weatherall and Boyer, 1962). The loci controlling the synthesis of $\alpha$, $\beta$, and $\gamma$ peptide chains are represented on the left. The arrows show the different dimers synthesized under the control of the two $\alpha$ alleles and of the $\beta$ and $\gamma$ genes. These dimers combine randomly in the last stage of assembly of the hemoglobin molecule to form the four hemoglobins indicated by capital letters.

$\alpha^A$ $\alpha^G$ heterozygous babies and the dimers combine randomly with $\beta_2^A$ or $\gamma_2^F$ dimers in a way apparently similar to that suggested to explain the presence of four hemoglobins in individuals doubly heterozygous for an $\alpha$ and a $\beta$ chain abnormality (Fig. 10).

### F. Hb-$A_2$ VARIANTS

#### 1. $\delta$ *Chain Variants*

Ceppellini (1959a) discovered an abnormal minor component, called Hb-$B_2$, which migrated in electrophoresis more slowly than Hb-$A_2$. Ceppellini (1959b) suggested that Hb-$B_2$ is related to Hb-$A_2$ since the amount of Hb-$A_2$ in the Hb-$B_2$ carriers is decreased to half its normal value and the sum of these two fractions equals the normal value of Hb-$A_2$. Moreover, in thalassemia heterozygotes the amount of both Hb-$A_2$ and Hb-$B_2$ is doubled (Ceppellini, 1959b).

Huisman and Meyering (1960) have also described a minor component with an electrophoretic mobility similar to that of Hb-$B_2$. Horton *et al.* (1961) have reported an individual apparently homozygous for the Hb-$B_2$ gene; the chemical study of Hb-$B_2$ has shown that only one peptide is different in the fingerprint of this hemoglobin, when compared to the fingerprint of Hb-$A_2$.

The peptide chains of Hb-$A_2$ and of Hb-$B_2$ have been separated and fingerprinted by Horton *et al.* (1961), and the altered peptide was shown to belong to the $\delta$ chain. The amino acid substitution may presumably involve a glycine in position 16 of the $\delta^{A_2}$ chain, which may be substituted by an arginine residue in the $\delta^{B_2}$ peptide chain (A. O. W. Stretton, personal communication); this suggestion is in accord with the chemical findings of Horton *et al.* (1961). These facts prove the existence of the $\delta$ gene as an independent genetic entity.

#### 2. *Formal Genetics of the* $\delta$ *Chain Variants*

Ceppellini (1959b) reported a double heterozygous propositus $\beta^A$ $\delta^{B_2}$/$\beta^S$ $\delta^{A_2}$ married to a normal woman. Their six children were found to be either Hb-S or Hb-$B_2$ carriers; the alternative segregation of the $\beta^S$ and $\delta^{B_2}$ genes suggests linkage of the $\beta$ and $\delta$ loci, the $\beta^S$ and $\delta^{B_2}$ genes being in repulsion in the propositus (see Fig. 11).

Huisman *et al.* (1961) have reported a family in which several members exhibit Hb-$B_2$ plus thalassemia minor of the "high-$A_2$" type (see Section IV,G). This type of thalassemia behaves as if allelic to the $\beta$ locus (see Ingram and Stretton, 1959b) and the abnormal gene may be indicated as $\beta^{Th}$ gene. The $\beta^{Th}$ and $\delta^{B_2}$ genes do not segregate in the family reported by Huisman *et al.* (1961), suggesting linkage of these two genes in coupling. It must be pointed out that in double heterozygotes for thalassemia and the $\delta^{B_2}$ gene, the increase of Hb-$A_2$ level is identical, whether the two mutant genes are in coupling $\beta^{Th}\,\delta^{B_2}/\beta^{A}\,\delta^{A_2}$ (Huisman *et al.*, 1961) or in repulsion $\beta^{Th}\,\delta^{A_2}/\beta^{A}\,\delta^{B_2}$ (Ceppellini, 1959a).

Not enough genetic data have been accumulated to determine the extent of the linkage between the $\beta$ and the $\delta$ loci. It seems unlikely, moreover, that enough data can be found to determine the fine genetic structure of the $\beta$ and $\delta$ loci. The $\beta$ and $\delta$ genes seem, however, to be under a common genetic control: (*1*) Hb-A and Hb-$A_2$ appear in fetuses at approximately the same age and increase in concentration in a parallel way (Kunkel *et al.*, 1957); (*2*) both Hb-A and Hb-$A_2$ are absent from individuals homozygous for the "hereditary persistence of Hb-F" (Wheeler and Krevans, 1961) or for the Hb-Pylos gene (Fessas *et al.*, 1962). These relationships in the behavior of the $\beta$ and $\delta$ genes and the family studies suggest a functional and topographical linkage of these genes.

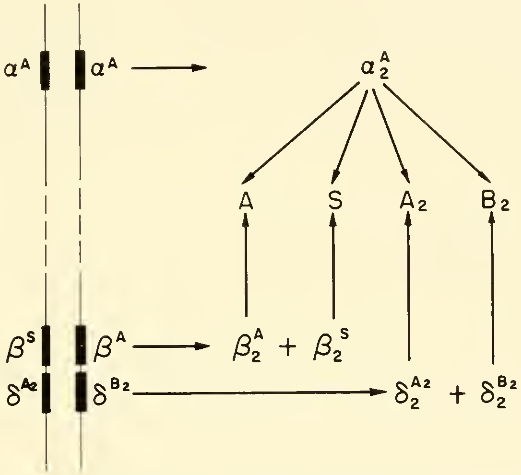


FIG. 11. Schematic explanation of the presence of four hemoglobins in the propositus of the family reported by Ceppellini (1959b). The genotype is illustrated on the left; the $\beta$ and $\delta$ loci are shown in linkage as suggested by the alternative segregation of the $\beta^S$ and $\delta^{B_2}$ genes. The arrows show the different dimers synthesized and their random combination to form the four hemoglobins indicated by capital letters.

### 3. *The $\alpha$ Chains of Hb-$A_2$*

Shooter *et al.* (1960) have shown in a person apparently homozygous for the abnormal hemoglobin $G_{\text{Ibadan}}$ the presence of a minor component with an altered electrophoretic mobility, called Hb-$G_2$. Hb-$G_{\text{Ibadan}}$ has been shown to have abnormal α chains by dissociation and recombination experiments (Gammack *et al.*, 1960) and it has been suggested that Hb-$G_2$ is a combination of $\alpha^{G_{\mathrm{Ibadan}}}$ chains with $\delta^{A_2}$ chains ($\alpha_2{}^{G_{\mathrm{Ibadan}}}\delta_2{}^{A_2}$). Recombination of Hb-$G_{\mathrm{Ibadan}}$ with Hb-$A_2$ produced some Hb-$G_2$ and Hb-A (Huehns and Shooter, 1961).

Several reports have appeared recently in the literature of a "split" Hb-$A_2$ in individuals who possess an abnormal hemoglobin. One of the minor components has the electrophoretic mobility characteristic of Hb-$A_2$, while the abnormal minor component shows the same difference in electrophoretic mobility relative to Hb-$A_2$ that the abnormal major component shows relative to Hb-A (Weatherall and Boyer, 1961). Boulard *et al.* (1961) have fingerprinted the abnormal minor component present in a Hb-I carrier and Baglioni (1962a) has fingerprinted the abnormal minor component present in a Hb-Norfolk carrier. Both Hb-I and Hb-Norfolk are known to have abnormal α chains (Murayama and Ingram, 1959; Baglioni, 1962a). The fingerprints of the abnormal minor components have shown the presence of α chains with alterations similar to the one reported for Hb-I and Hb-Norfolk, respectively, and of normal $\delta^{A_2}$ chains.

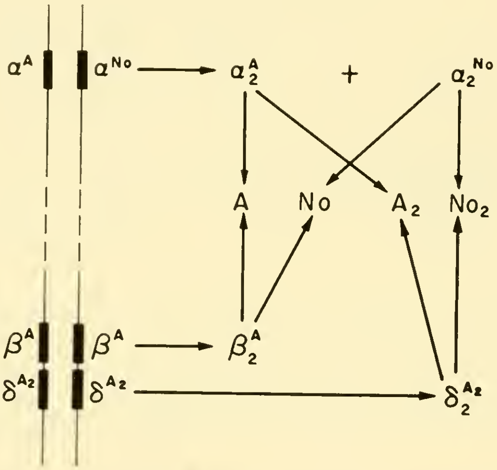


FIG. 12. Schematic explanation of the presence of four hemoglobins in individuals heterozygous for an α chain mutation (Baglioni, 1962a). The genotype is illustrated on the left (No = Norfolk). The arrows show the different dimers synthesized under the control of the two α alleles and of the β and δ genes. The dimers combine randomly to form the four hemoglobins indicated by capital letters. The β and δ loci are linked (Ceppellini, 1959b).

The α chains of Hb-A and of Hb-$A_2$ have been shown to be chemically identical (Ingram and Stretton, 1959; Muller and Jonxis, 1960); they are under the control of the same $\alpha$ gene. The $\delta_2$ dimers, like the $\gamma_2$ dimers, combine in a random way with all the types of $\alpha_2$ dimers present in the last stage of Hb-$A_2$ synthesis (Fig. 12).

### 4. *The Problem of Hb-Lepore*

Gerald and Diamond (1958) described a patient of Italian extraction affected by a severe anemia, clinically classified as thalassemia major. The propositus, his mother, and four maternal relatives showed an electrophoretically slow moving abnormal hemoglobin called Hb-Lepore, from the name of the family, in concentration of approximately 10 to 15%; they were found to be affected by microcytemic anemia. The father of the propositus was found to be a $\beta^{Th}/\beta^{A}$ heterozygote and the severe anemia in the propositus was thought to result from the interaction of the $\beta^{Th}$ gene and of the Hb-Lepore gene. Similar cases of an electrophoretically slow moving abnormal hemoglobin, associated with microcytemic anemia have been reported among Italians (Silvestroni and Bianco, 1958), Papuans (Neeb *et al.*, 1961), and Greeks (Fessas *et al.*, 1962). This condition seems to be relatively frequent among patients with microcytemic anemia in Greece (Fessas *et al.*, 1962).

The abnormal hemoglobin, always present in concentration of about 10–15%, has been alternatively classified as Hb-G (Silvestroni and Bianco, 1958), Hb-Pylos (Fessas *et al.*, 1962), or Hb-Lepore (Neeb *et al.*, 1961). Individuals apparently homozygous for the Hb-Lepore gene or for the Hb-Pylos gene have been reported; they are characterized by a complete absence of Hb-A and of Hb-$A_2$ in the case described by Fessas *et al.* (1962) and presumably in the two cases of Neeb *et al.* (1961). In the heterozygotes the amount of Hb-$A_2$ is reduced to about half the normal value (Fessas *et al.*, 1962).

Gerald *et al.* (1961) have briefly reported the results of their chemical studies on Hb-Lepore and Hb-Pylos; these hemoglobins give fingerprints indistinguishable from those of Hb-$A_2$. Moreover, some of the peptides of Hb-Lepore were analyzed and found to have an amino acid composition identical to homologous peptides of Hb-$A_2$. Neeb *et al.* (1961) have also fingerprinted Hb-Lepore from Papuans and have obtained a peptide pattern indistinguishable from that of Hb-$A_2$. Hb-Lepore has presumably normal $\alpha$ peptide chains, since no alteration of the fingerprint pattern of the $\alpha$ chain has been observed, but mainly because normal Hb-F is present in Hb-Lepore carriers; one would expect that if any abnormal $\alpha_2$ dimers were present, they would associate randomly with $\gamma_2$ as well as $\delta_2$ dimers. The abnormality in the "$\delta$" chain of Hb-Lepore has not been discovered, but it may possibly be hidden in the "core" of this hemoglobin, which is excluded from the fingerprinting analysis. The Hb-Lepore abnormality thus seems to involve a mutation of the $\beta$ locus, resulting in complete suppression of $\beta$ chain synthesis and a simultaneous mutation of the $\delta$ locus, resulting in the synthesis of an electrophoretically abnormal hemoglobin, Hb-Lepore. It is rather hard to understand why these independent mutations in the two linked genes $\beta$ and $\delta$ should always be observed in coupling. The $\delta^{B_2}$ mutation has, for instance, been observed either in coupling or in repulsion with a $\beta^{Th}$ gene (Ceppellini, 1959a; Horton *et al.*, 1961).

It seems likely that a common genetic event is responsible for both the absence of $\beta$ chains and the alteration of the "$\delta$" chains; the only genetic event involving two neighboring genes at the same time is an overlapping deletion of genetic material, resulting either from mutation or from unequal crossing-over.

It has recently been suggested by Smithies *et al.* (1961) that a mechanism of unequal crossing-over between two haptoglobin genes could be responsible for the duplication in length of one peptide chain of the human haptoglobins. A similar mechanism can be advocated to explain the puzzling genetic and chemical findings on Hb-Lepore: in this case loss of genetic material by unequal crossing-over, may result in the formation of a hybrid $\beta$-$\delta$ gene, having part of the genetic material of both the $\beta$ and the $\delta$ genes.

The chemical findings on Hb-Lepore so far do not contradict this theory; only that part of the "$\delta$" chain of Hb-Lepore, going from the N-terminal end of the peptide chain to approximately halfway along the chain, has been shown to give a fingerprint similar to the $\delta$ chain of Hb-$A_2$, while no information exists about the C-terminal portion of the "$\delta$" chain of Hb-Lepore. The possibility exists that the C-terminal part of the "$\delta$" chain of Hb-Lepore is identical to the $\beta$ chain of Hb-A.[3]

It is known that unequal crossing-over frequently involves chromosomal duplications like the *bar* duplication of *Drosophila melanogaster* (Bridges, 1936); *double-bar* and normal flies appear with relatively high frequency as a result of unequal crossing-over in homozygous *bar* females.

It has been suggested that the $\delta$ gene results from a duplication of the $\beta$ gene, followed by independent evolution (Ingram, 1961a). Profound homologies exist between the $\beta$ gene and the $\delta$ gene, since the primary sequence of these peptide chains is extremely similar (Ingram and Stretton, 1961). These circumstances may increase the probability of unequal crossing-over involving the $\beta$ and $\delta$ loci.

---

[3] Recent experiments (Baglioni, unpublished) have shown that the C-terminal portion of the $\delta$ chain of Hb-Lepore is identical to the C-terminal portion of the $\beta$ chain of Hb-A. Two amino acid replacements have been located by A. O. W. Stretton (unpublished) in the C-terminal portion of the $\delta$ chain of Hb-$A_2$. These replacements are not present in the $\beta$-$\delta$ chain of the Hb-Lepore.

### G. THE PROBLEM OF THALASSEMIA

*Thalassemia* is a hereditary anemia which occurs commonly in Mediterranean countries. Two main forms of the disease are recognized: *thalassemia minor*, which occurs in individuals heterozygous for a thalassemic gene, and *thalassemia major* or Cooley's anemia, in homozygotes (Valentine and Neel, 1944). *Thalassemia minor* is a mild hypochromic microcytic anemia which may be asymptomatic, while *thalassemia major* is a marked hypochromic anemia, characterized by anisocytosis and accompanied by profound alterations of skull and bones, due to bone marrow hyperplasia. Children affected by thalassemia major usually do not reach adulthood and early death in the absence of constant medical care and repeated blood transfusions is the rule. Several aspects of the thalassemic diseases have recently been reviewed by Lehmann (1960) and by Rucknagel and Neel (1961).

#### 1. *Hemoglobins in Thalassemia*

In thalassemia major Hb-F is constantly present in amounts which, together with the hematological picture, are characteristic of this disease. As much as 97% of the total hemoglobin has been reported to be Hb-F in a thalassemic patient (Rich, 1952). This Hb-F appears to be identical to Hb-F isolated from cord blood when analyzed by fingerprinting (Baglioni, unpublished). In thalassemia minor the proportion of Hb-F is very close to normal values.

Kunkel and Wallenius (1955), and subsequently Ceppellini (1956) and several others, have reported an increase of Hb-$A_2$ to values approximately double the normal values in individuals affected by thalassemia minor. The Hb-$A_2$ obtained from thalassemic patients has been found to be indistinguishable from Hb-$A_2$ from normal subjects by fingerprinting (Ingram and Stretton, 1961). Kunkel *et al.* (1957) and Carcassi *et al.* (1957) found several exceptions to the rule of increased Hb-$A_2$ in thalassemia minor. This and other genetic observations have proved the heterogeneity at the molecular level of thalassemia.

Several combinations of a thalassemic gene with abnormal hemoglobin genes have been reported in the literature (see Lehmann, 1960). Affected individuals in general show a moderately severe microcytic anemia; the range of variability is, however, quite wide. Sickle-cell/thalassemia is the combination which is more frequently observed (Silvestroni and Bianco, 1952). Hb-S, Hb-F, and very small amounts of Hb-A may be present in sickle-cell/thalassemia. Since the amount of any abnormal hemoglobin is found to be lower than that of Hb-A in heterozygotes for an abnormal hemoglobin gene, the inversion of the Hb-A/abnormal hemoglobin ratio has been thought to be characteristic of thalassemia in combination with an abnormal hemoglobin gene. This type of thalassemia has been defined "interacting" thalassemia and combinations of this type of thalassemia with Hb-C, Hb-D, and Hb-E have also been described (quoted in Lehmann, 1960). These hemoglobins have been shown to have altered $\beta$ chains (Hb-D is presumably of the Hb-$D_{Punjab}$ type).

Thalassemic individuals in whom the ratio Hb-A/abnormal hemoglobin is not inverted have been reported for the Hb-S/thalassemia combination and for the Hb-C/thalassemia combination (Zuelzer *et al.*, 1956; Zuelzer and Kaplan, 1954). The thalassemia in these individuals is thus of a "non-interacting" type. Examination of some of these individuals has shown that the non-interacting thalassemia is of the non-increased Hb-$A_2$ type. Interaction of the non-increased Hb-$A_2$ type thalassemia with Hb-I, resulting in an inversion of the Hb-A/Hb-I ratio has been reported by Atwater *et al.* (1960b). Hb-I is known to be altered in the $\alpha$ chain (Murayama and Ingram, 1959). Thalassemic genes may thus interact with $\alpha$ or $\beta$ mutant genes and be of the non-increased or increased Hb-$A_2$ type, respectively.

### 2. *Formal Genetics of Thalassemia*

Several marriages of persons affected by Hb-S/thalassemia or Hb-C/thalassemia combinations have been reported in the literature. (For a review, see Rucknagel and Neel, 1961.) In most marriages only children either heterozygous for thalassemia or heterozygous for the abnormal S or C hemoglobins are born. The thalassemia in these families is of the increased Hb-$A_2$ interacting type. The thalassemic gene in these families behaves as if allelic to the $\beta^S$ or $\beta^C$ genes. However, in some similar marriages, normal children, or children with thalassemia or sickle-cell trait or with sickle-cell/thalassemia, are born (quoted in Ingram and Stretton, 1959b). These findings indicate non-allelism of the thalassemic gene concerned with the $\beta^S$ or $\beta^C$ genes. The thalassemia in these families is usually of the non-increased Hb-$A_2$, non-interacting type.

In the marriage of the Hb-I thalassemic individual reported by Atwater *et al.* (1960b), only children with the thalassemic trait were born. It may be pointed out that this thalassemic gene is of the non-increased Hb-$A_2$ type and was shown to interact with the $\alpha^I$ gene; it is not proved, because of the very small number of individuals in this family, whether this thalassemic gene behaves as if allelic to the $\alpha^I$ gene, as the family data suggest.

### 3. *The Molecular Basis of Thalassemia*

The main pathological manifestation of thalassemia is the failure to produce red cells with a normal hemoglobin content. The total red cells production is increased (Bailey and Prankerd, 1958), but in a way insufficient to compensate for the decreased hemoglobin content of the red cells.

The decrease in hemoglobin synthesis is at the expense of Hb-A only: abnormal hemoglobins, found in heterozygous combination with thalassemia, are produced at a normal rate, so that defects of the heme synthesis or of other enzymatic mechanisms seem unlikely in these cases.

Abnormal hemoglobins are usually found in lower amounts than Hb-A in heterozygotes for an abnormal hemoglobin gene. It is assumed that the lower rate of synthesis of abnormal hemoglobins relative to Hb-A is the consequence of the gene mutation. The ratio Hb-A/abnormal hemoglobin is, however, quite variable, from close to 1 to 4 or more (see Section V,A,2). One can thus imagine instances where a hemoglobin gene mutation produces such drastic alterations that the corresponding abnormal hemoglobin cannot be made at all or is made at a very low rate.

Itano (1956), and in more detail Ingram and Stretton (1959b), have suggested that the drastic reduction in Hb-A synthesis in thalassemia may result from an alteration or an amino acid substitution (Ingram and Stretton, 1959b), not involving a change of the electrophoretic charge of the hemoglobin molecule.

Either peptide chain of Hb-A could be involved by such alteration and two types of thalassemia can be defined (Ingram and Stretton, 1959b): (*a*) $\alpha$ chain thalassemia, allelic to the $\alpha$ locus, which impairs the synthesis of $\alpha$ chains; (*b*) $\beta$ chain thalassemia, allelic to the $\beta$ locus, which impairs the synthesis of $\beta$ chains. This hypothesis provides an explanation for most of the alterations of the hemoglobin pattern in thalassemia. In the $\alpha$ chain thalassemia a reduced amount of $\alpha_2$ dimers is available to combine with $\beta_2^A$ and $\delta_2^{A_2}$ dimers or with any abnormal type of $\beta_2$ dimers, like $\beta_2^S$ or $\beta_2^C$. Consequently, the rate of synthesis of Hb-A, Hb-$A_2$, and of any other hemoglobin containing $\alpha_2$ dimers will be equally lowered and the relative proportion of these hemoglobins will be unchanged from the normal; the $\alpha$ chain thalassemia is in fact of a non-increased Hb-$A_2$ and non-interacting with $\beta$ chain abnormality type. On the contrary, the $\alpha$ chain thalassemia will interact with $\alpha$ chain abnormal genes, like $\alpha^I$ (Atwater *et al.*, 1960b). The balance in the rate of synthesis of $\alpha$ chain versus $\beta$, $\gamma$, and $\delta$ chains is profoundly altered in $\alpha$ thalassemia; $\beta_2$, $\gamma_2$, and $\delta_2$ dimers are produced in large excess over the available $\alpha_2$ dimers. Aggregation of identical dimers may in this case result in the formation of abnormal hemoglobin molecules, like Hb-H ($\beta_4^A$) and Hb-Bart's ($\gamma_4^F$) (see Section IV,H). In $\beta$ chain thalassemia there is a deficit in the $\beta$ chain production; in the heterozygote $\beta^A/\beta^{Th}$, for instance, only the $\beta^A$ gene is operating and the rate of synthesis of $\beta$ chains is lowered to approximately half the normal value. Since the rate of synthesis of $\delta$ chains is not affected by the presence of a thalassemic gene, the relative amount of Hb-$A_2$ appears to be doubled over normal values. In sickle-cell/thalassemia and in other combinations of a thalassemic gene with one of the $\beta$ alleles, such as $\beta^C$, $\beta^E$, or $\beta^D$, the rate of production of $\beta^C$, $\beta^E$, or $\beta^D$ chains, respectively, is higher than that of the $\beta$ chains produced under the control of the $\beta^{Th}$ gene. $\beta$ chain thalassemia is thus of an increased Hb-$A_2$ type and interacts with $\beta$ alleles.

The denomination thalassemia includes a group of hereditary anemias heterogeneous from a genetic point of view; thalassemia minor may be caused by heterozygosity for several possible $\alpha^{Th}$ or $\beta^{Th}$ pseudoalleles and thalassemia major by homozygosity for two $\alpha^{Th}$ or $\beta^{Th}$ genes or possibly by interaction of one $\alpha^{Th}$ with one $\beta^{Th}$ gene. The heterogeneity of thalassemia is in accordance with the variability of the hematological and clinical picture; the possibility exists that different thalassemic genes have different expressivity and penetrance. A form of thalassemia of intermediate severity, which had been designated *thalassemia intermedia*, may possibly result from the combination of thalassemic genes with low expressivity; other factors besides the hemoglobin genotype may also control the expression of the thalassemic genes. The phenotypic effects of the thalassemic genes are consistent with the hypothesis of the amino acid substitution involving non-charged amino acids (Ingram and Stretton, 1959b). However, no chemical alteration of the hemoglobin of thalassemic patients has been reported so far. Few research groups have been involved in the search for the hidden amino acid substitution in thalassemia, and the author is among them. In spite of extensive examination of the "Hb-A" from patients with thalassemia major, no alteration of the fingerprinting pattern of this hemoglobin, compared to Hb-A from normal subjects, has yet been observed (C. Baglioni, V. M. Ingram, and A. O. W. Stretton, unpublished). The analysis of the peptides of the "Hb-A" isolated from a sickle-cell/thalassemic patient has not revealed any difference in amino acid composition with peptides isolated from authentic Hb-A (G. Guidotti, personal communication). The evidence against the amino acid substitution hypothesis is, however, not conclusive at all; more work and more chemical data are necessary before this possibility can be excluded. Inversions of the amino acid sequence, for instance, are probably not revealed by the fingerprinting analysis or by the analysis of the amino acid composition of peptides; the only conclusive chemical evidence is the determination of the entire primary sequence of the "Hb-A" from a thalassemic. To reach this goal may require years of hard work.

Moreover, very little is known about the effect of a change in the genetic information of a structural gene on the primary sequence of the corresponding peptide chain. Some mutations may not result in an alteration of the primary structure of a peptide chain, but only in a difficult "reading" of the genetic code: i.e., ambiguity in the genetic information may cause the production of a large proportion of senseless peptide chains unable to fold, or a difficulty in translating the genetic code may slow down the synthesis of a peptide chain. An alternative hypothesis has been advanced by Ingram and Stretton (1959b), based on Freese's (1958) idea of "connecting units" alternating in linear arrangement with structural genes. If these connecting units serve to induce or repress structural genes, then thalassemia may be the consequence of a change of these regulators. However, thalassemia is not likely to be a mutation of an "*i*" type gene, in the terminology of Jacob and Monod (1961); mutations of these genes exhibit a *trans* effect, which is not observed in thalassemia. Only the synthesis of the $\alpha$ or $\beta$ peptide chain in *cis* with the thalassemic mutation is repressed. The validity of a comparison between the function of the genes of a bacterial chromosome with the genes of higher organisms' chromosomes is, however, questionable. The complexity of chromosomal structures in higher organisms and the larger dimensions of the genetic material may limit the diffusion of repressor or inducer molecules to neighboring groups of genes and mimic in this way a *cis* effect.

### H. ABNORMAL HEMOGLOBINS FROM ALTERED AGGREGATION OF PEPTIDE CHAINS

Under this denomination are classified abnormal hemoglobins which are not properly described as variants of the normal hemoglobins.

#### 1. *Hb-H*

Rigas *et al.* (1955) and Gouttas *et al.* (1955) described this electrophoretically fast moving hemoglobin, characterized by instability and denaturation on storage. Hb-H is precipitated in the erythrocytes as inclusion bodies on incubation with cresyl blue.

Jones *et al.* (1959a) have shown that Hb-H is made up of four $\beta^A$ chains. Motulsky (1956) pointed out the close association between Hb-H and thalassemia; Hb-H carriers usually present the blood picture characteristic of thalassemia and one parent is found to be affected by thalassemia minor. The thalassemia in the Hb-H carrier and in the affected parent is of the non-increased Hb-$A_2$ type (Silvestroni *et al.*, in the infants, when Hb-A substitutes Hb-F. The presence of Hb-Bart's in cord blood has been attributed in these cases to heterozygosity for $\alpha$ chain thalassemic genes. These genes may manifest themselves with ever, of no use for the carriers; its oxygen affinity is far too high to permit release of oxygen to the tissues (Benesch *et al.*, 1961).

Hb-H is not inherited in a Mendelian way (Motulsky, 1956). The parent affected by thalassemia does not show, in the majority of the cases reported in the literature, any Hb-H; transmission from father to son has been reported (Fessas, 1959). Only a small fraction of the individuals presumably affected by $\alpha$ thalassemia are Hb-H carriers. The unpredictable appearance of Hb-H may be related to other factors, which affect the erythropoiesis or the rate of destruction of red cells in subjects affected by $\alpha$ thalassemia. The amount of Hb-H is quite variable in the carriers and undetected amounts of this hemoglobin may be present in individuals classified as normal. The presence of Hb-H may also be caused by the interaction of two abnormal genes, one of which has a very low expressivity and does not appreciably depress hemoglobin synthesis in heterozygotes.

### 2. *Hb-Bart's*

Fessas and Papaspyrou (1957) and Ager and Lehmann (1958) reported an abnormal hemoglobin in infants with a thalassemic blood picture. This hemoglobin has been designated Hb-Bart's. Hunt and Lehmann (1959) have fingerprinted Hb-Bart's and have shown that this hemoglobin is made up of $\gamma^F$ chains only; Hb-Bart's seems to have the structural formula $\gamma_4^F$. Like Hb-H, Hb-Bart's has been found to be associated with $\alpha$ thalassemia. Baglioni *et al.* (1961) have shown the presence of Hb-Bart's in a newborn, who inherited a thalassemic gene interacting in his mother with an $\alpha$ chain abnormality ($\text{Hb-I} = \alpha_2^I \beta_2^A$) and thought to be an $\alpha^{Th}$ gene. Recent studies on the hemoglobins of newborns have shown that Hb-Bart's occurs with high frequency among newborns in some populations. Vella (1959) observed a fast abnormal component, presumably identical with Hb-Bart's, in 4% of the cord bloods of Chinese newborns. Hendrickse *et al.* (1960) have found a 10% frequency of Hb-Bart's in cord bloods of Nigerian newborns. Similar findings for the Negro population of the United States have been reported by Minnich *et al.* (1962) and by Weatherall and Boyer (1961). Lie-Injo (1961) has reported the occurrence of Hb-Bart's in Indonesians, Malays, and Chinese. The high frequence with which Hb-Bart's has been observed in Negroes is quite surprising. Hb-Bart's disappears almost at the same time that Hb-F disappears and usually no Hb-H is observed in the infants, when Hb-A substitutes Hb-F. The presence of Hb-Bart's in cord blood has been attributed in these cases to heterozygosity for $\alpha$ chain thalassemic genes. These genes may manifest themselves with the presence of Hb-Bart's only in periods of stress for the erythropoietic system, when the production of red cells is accelerated by environmental or physiological factors around birth. The frequency with which Hb-Bart's is observed in newborns would correspond to the true frequency of such genes in population, unaffected by expressivity or penetrance factors.

If the frequency of $\alpha^{\mathrm{Th}}$ genes is as high as estimated by the above-mentioned authors, one wonders why homozygous cases of $\alpha$-thalassemia have not been observed and reported in the literature. An answer to this question may be found in the description by Lie-Injo (1961) of several cases of severe hydrops and erythroblastosis fetalis in stillborn babies or in babies who died shortly after birth. The hematological picture was that of a severe hemolytic disease and the red cells showed sickling although no Hb-S was present. The hemoglobin of these stillborn babies consisted mainly of Hb-Bart's with small amounts of Hb-F and Hb-A (Lie-Injo and Lie, 1961). In a cord blood sample examined by the author (obtained through the courtesy of Dr. Lie-Injo, Institute for Medical Research, Kuala Lumpur, Malaya) approximately 80% of the hemoglobin was found to be Hb-Bart's! Lie-Injo (personal communication) has suggested that these babies were affected by homozygous $\alpha$-thalassemia. The failure to observe cases of homozygous $\alpha$-thalassemia in adults may thus be due to the fetal or stillborn lethality of this genotype. Since $\alpha$ chains are common to Hb-F, Hb-A, and Hb-$A_2$, the synthesis of these hemoglobins will be equally affected by a deficiency of $\alpha$ chain production. Moreover, no compensatory effect is possible in homozygous $\alpha$ thalassemia, while in $\beta$ thalassemia the production of high quantities of Hb-F allows the affected individuals to live for a few to several years.

It is interesting that no abnormal aggregation of $\alpha$ chains to form a hypothetical $\alpha_4$ hemoglobin is observed in $\alpha$ thalassemia. The $\alpha_4$ molecule does not form either because of steric hindrance or because of a lack of chemical affinity of the $\alpha$ chains to polymerize. It has indeed been observed that upon prolonged exposure of Hb-A to acid pH and recombination, small quantities of $\beta^{\mathrm{A}}$ chains polymerize to form Hb-H and that an equivalent amount of $\alpha^{\mathrm{A}}$ chains are liberated as monomers and not as $\alpha_4{}^{\mathrm{A}}$ tetramers (Huehns *et al.*, 1961b); it is not known whether these $\alpha$ chains are altered or partially denaturated.

### 3. *Hb-Augusta I and Hb-Augusta II*

Huisman (1960) has reported the occurrence of a fast minor component in the cord blood of Negro babies, together with Hb-F, Hb-A, and Hb-S. Dissociation and recombination analysis of this minor component, called Hb-Augusta I, showed the presence of $\beta^S$ chains only and the structural formula $\beta_4^S$ was assigned to this hemoglobin. Family studies showed the presence of minute amounts of Hb-H in one parent and in some of the brothers of the propositus; it was concluded that the formation of Hb-Augusta I ($\beta_4^S$) was caused by a genotype, $\alpha^{Th}/\alpha^A\ \beta^S/\beta^A$.

Preliminary evidence has been reported by Huisman (1960) that a different minor component observed in the cord blood of a baby, presumably heterozygous for the Hb-C gene and for a thalassemic gene ($\alpha^{Th}/\alpha^A\ \beta^C/\beta^A$), is made up of $\beta^C$ chains only; this $\beta_4^C$ hemoglobin has been designated Hb-Augusta II.

## V. The Control of the Synthesis of Hemoglobin

### A. THE CONTROL OF THE RATE OF SYNTHESIS OF THE PEPTIDE CHAINS

### 1. *Rate of Synthesis of Normal Peptide Chains*

Approximately 95% of the red cell protein is hemoglobin. The peptide chains of hemoglobins are thus synthesized at a rate much higher than that of any other protein in the red cells and the hemoglobin synthesis is one of the main aspects of the differentiation of stem cells into erythrocytes. Very little is known of the mechanisms controlling this differentiation. One can guess that in the course of differentiation the hemoglobin genes become activated in such a way that their products become the main components of the red cells; however, how the activation of a group of genes is brought about selectively in mammalian tissues is not known.

In the normal red cells only "complete" hemoglobin molecules are observed, i.e., Hb-A $= \alpha_2^A \beta_2^A$, Hb-F $= \alpha_2^A\ \gamma_2^F$, or Hb-A$_2$ $= \alpha_2^A\ \delta_2^{A_2}$. These hemoglobin molecules have in common $\alpha_2^A$ dimers. During hemoglobin synthesis the amount of $\beta_2^A$, $\gamma_2^F$, and $\delta_2^{A_2}$ dimers produced is equivalent to the amount of $\alpha_2$ dimers produced. This raises the question of how this equilibrium in the synthesis of products of independent genes is maintained. It is known that when a defective $\alpha$ gene is present (as in $\alpha$ chain thalassemia), abnormal aggregation of $\beta_2$ or $\gamma_2$ dimers may be observed. It may thus be supposed that the $\beta$ and $\gamma$ genes have an autonomous regulation of $\beta$ and $\gamma$ chain production. In pathological conditions in which an equivalent amount of $\alpha$ chains is not produced, the $\beta$ genes during adult life or the $\gamma$ genes during fetal life produce $\beta_2$ and $\gamma_2$ dimers in excess over the available $\alpha_2$ dimers. A similar situation is not observed in conditions involving a decrease in $\beta$ chains production (as in $\beta$ thalassemia). No abnormal aggregation of $\alpha_2$ dimers is observed in this disease and no evidence of an accumulation of these dimers has been reported. The $\alpha$ gene seem thus to have a non-autonomous regulation. The amount of $\alpha$ chains produced seems to be dependent on the amount of $\gamma$ or $\beta$ chains available; if not enough $\beta$ or $\gamma$ chains are made, the production of $\alpha$ chains seems to be lowered in a parallel and corresponding way. It is not known how this control is achieved; one can guess that steric hindrance factors, prohibiting the polymerization of $\alpha$ chains, are of importance, if the solubility of the dimers or of the single chains is low in physiological conditions. One can visualize an equilibrium between $\alpha$ peptide chains sitting on or in the proximity of the ribosomal templates, dimerization of the $\alpha$ chains, and release of the $\alpha_2$ dimers upon combination with $\beta_2$, $\gamma_2$, or $\delta_2$ dimers. The answer to these questions may possibly come from experiments on hemoglobin synthesis in reticulocytes or in cell-free systems (Dintzis, 1961) (see Schweet and Bishop, Chapter VIII).

Within the hemoglobin peptide chains there are extreme variations in the rate of synthesis. Hb-A$_2$ ($\alpha_2{}^{A}\ \delta_2{}^{A_2}$) represents approximately only 2.5% of the hemoglobin in normal adults. For every 40 $\beta^{A}$ chains only one $\delta^{A_2}$ chain is synthesized. This ratio is quite constant, within the limits of random variability, and is only found to be altered in favor of the $\delta$ chains when $\beta$ chains are produced with lower efficiency, as in $\beta$ thalassemia. One wonders how this constant ratio is maintained and why the $\delta$ chains, which are extremely similar to the $\beta$ chains from a chemical point of view, are produced with so much lower an efficiency. The existence of controlling genetic elements has been postulated by several authors (see Neel, 1961); in their view the synthesis of the $\delta$ chains is repressed by a specific regulator gene or, vice versa, the synthesis of $\beta$ chains is activated to a higher extent by a specific regulator gene. Thalassemia may then correspond to a mutation of a regulator gene, rather than of a structural gene (Ingram and Stretton, 1959b). No clear-cut genetic or chemical evidence has so far been presented in support of this hypothesis.

### 2. *The Rate of Synthesis of Abnormal Peptide Chains*

The finding of Pauling *et al.* (1949) that an abnormal hemoglobin is present in lower amounts than Hb-A has been confirmed for all the abnormal hemoglobins discovered. The only possible exception to this rule reported in the literature is Hb-J (Thorup *et al.*, 1956).

The rate of synthesis of an abnormal peptide chain is, in general, lower than that of the corresponding normal chain. The Hb-A/abnormal hemoglobin ratio, which is assumed to be equivalent to the ratio normal chain/abnormal chain, is quite variable for different abnormal hemoglobins and for different individuals. Itano (1953) has reported the distribution of Hb-S and Hb-C in members of several families, pointing out the variability of the ratio Hb-A/Hb-S or Hb-A/Hb-C, and introducing the concept of relative rates of synthesis as the determining factor in hemoglobin ratios. Itano (1953) suggests that the variability in the ratio Hb-A/abnormal hemoglobin may be dependent on the relative ability to synthesize Hb-A, rather than on the different rate of synthesis of the abnormal hemoglobin. Chernoff (1956) has reported the distribution of Hb-S, Hb-C, and Hb-D in heterozygotes, observing conspicuous variations in the amount of these hemoglobins relative to Hb-A. Several genetic and environmental factors are likely to determine this variability; no definite correlation has, however, been established between the Hb-A/abnormal hemoglobin ratio and any specific factor.

Different abnormal hemoglobins are present in different amounts relative to Hb-A. On examining data reported in the literature one observes a fairly wide range of variation; it may thus be supposed that the abnormal peptide chains are synthesized at a characteristic rate, determined by the specific alteration of these peptide chains. How does a single amino acid substitution affect the rate of synthesis of peptide chains 141–146 amino acids in length? At which stage in the synthesis of a peptide chain does the amino acid substitution become a rate-limiting factor? It does not seem likely that an amino acid substitution may decrease the rate at which the amino acids are aligned along the ribosomal template and bound in peptide linkage for the synthesis of a peptide chain. If this were the case, abnormal hemoglobins having the same type of amino acid substitution should have an identical rate of synthesis; this does not seem to happen for Hb-C, Hb-E, and Hb-$0_{\mathrm{Arabia}}$, which have a glutamic acid substituted by a lysine in different positions of the $\beta$ chain. The ratio Hb-A/abnormal hemoglobin in heterozygotes for Hb-C, Hb-E, and Hb-$0_{\mathrm{Arabia}}$ has been found to be different.

The following steps in the synthesis of the hemoglobin molecule are the folding of the individual peptide chains, the dimerization of identical chains, and the association of symmetric dimers. As far as we know, any of these processes can be disturbed by the presence of a wrong amino acid in a peptide chain; however, the most likely to be disturbed seems to be the folding of a peptide chain to assume its secondary or tertiary configuration.

### B. THE CONTROL OF THE SYNTHESIS OF FETAL AND ADULT HEMOGLOBIN

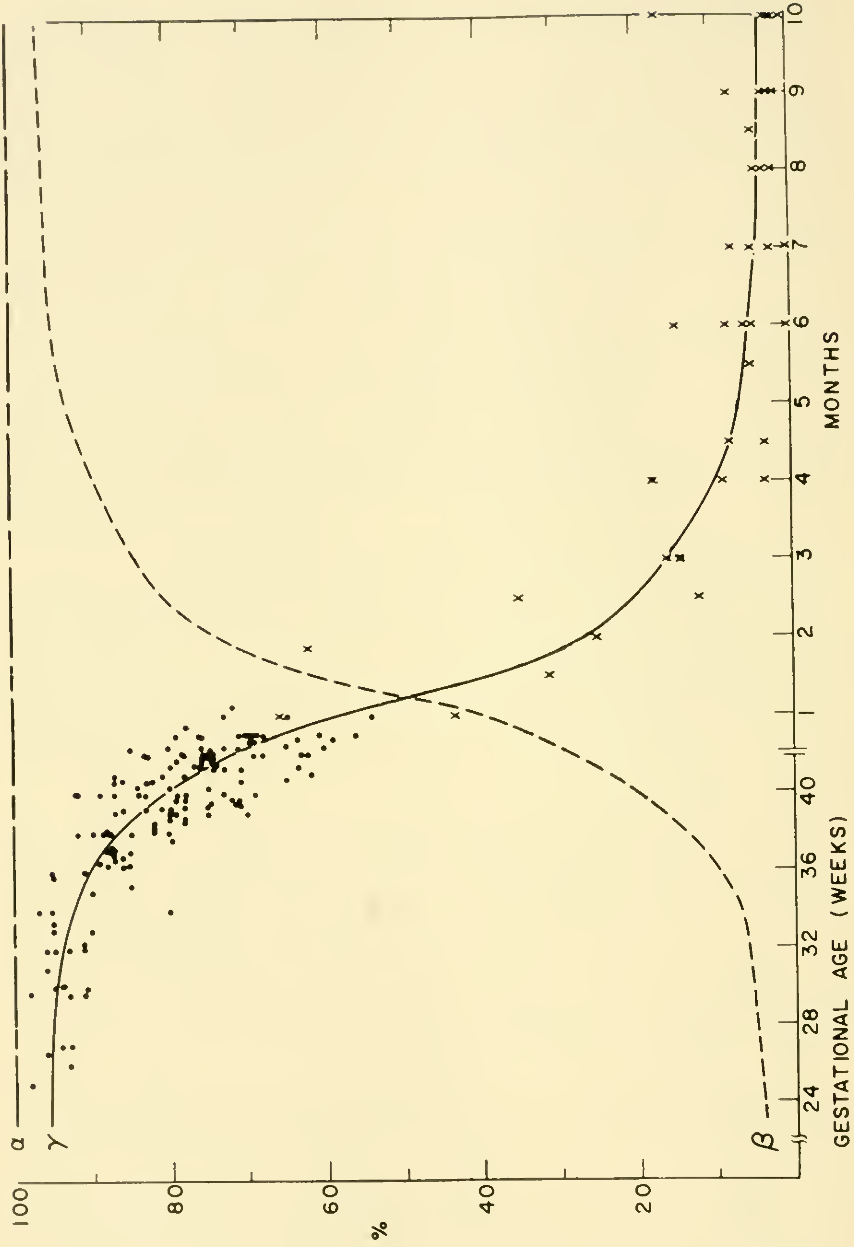
Hb-F is the principal component of human fetus hemoglobin. The synthesis of Hb-A starts, however, early in fetal life; Hb-A has been detected in fetuses of 13 weeks (Walker and Turnbull, 1955). The level of Hb-A raises to about 10% by 22 weeks and to an extremely variable 20% at birth. The Hb-F level is fairly constant from the twenty-fourth to the thirty-second week of fetal life and then falls at a rate of 2.5 to 4% per week, Hb-F being substituted by Hb-A (Cook *et al.*, 1957).

The level of Hb-F at birth in infants of different gestational ages has been determined by Cook *et al.* (1957); the concentration of Hb-F decreases with increasing gestational age of the newborns (Fig. 13). Babies born prematurely have higher levels of Hb-F, while babies born after term have lower levels of Hb-F; i.e., babies born 4 weeks after term seem to have Hb-F levels comparable to those of 4-week-old babies born at the normal term. Betke (1958) has also studied the fall of Hb-F concentration in newborns and has come to the conclusion that the Hb-F level falls as a function of gestational age and that birth has a negligible effect upon the variation of Hb-F level.

Hb-F is present for a variable time after birth (Chernoff and Singer, 1952; Jonxis and Visser, 1956); the Hb-F level decreases to about 1% or less by the end of the first year of adult life. The disappearance of Hb-F takes significantly longer than the time predicted on the basis of a complete cessation of Hb-F synthesis (White and Beaven, 1959); the production of Hb-F decreases after birth and is very limited in normal adults.

In the course of fetal differentiation, erythropoietic activity resides in different morphological sites. Hemoglobin containing cells appear first in the mesenchyme of the embryo; the liver and the spleen are the next active erythropoietic sites until birth (Maximow and Bloom, 1957). The bone marrow is the main center of red cell production around birth; after birth, active bone marrow disappears from the long bones and the erythropoietic activity is limited to the red marrow of the sternum and vertebrae (Maximow and Bloom, 1957).

No correlation has been established between the morphological sites of erythropoiesis and the switch from Hb-F synthesis to Hb-A synthesis. Betke (1958) has shown that Hb-F and Hb-A are both present in red cells obtained from liver and bone marrow of human fetuses. Thomas *et al.* (1960) have determined the hemoglobin types synthesized by fetal spleen, liver, and bone marrow; these hematopoietic tissues synthesize both Hb-A and Hb-F, although in a different ratio. More Hb-A is synthesized in bone marrow preparations than in spleen or liver prepara-


α
γ
β
%
100
80
60
40
20
24
28
32
36
40
GESTATIONAL AGE (WEEKS)
1
2
3
4
5
6
7
8
9
10
MONTHS

tions; bone marrow is the latest hematopoietic site in fetuses and red cells produced in bone marrow are presumably more differentiated toward the Hb-A containing red cell type.

Kleihauer *et al.* (1957) have devised a method to demonstrate Hb-F within single cells, by washing blood smears with buffer solutions, which selectively dissolve Hb-A; the hemoglobin left in the red cells is assumed to be Hb-F. By this method of analysis Kleihauer *et al.* (1957) have shown that Hb-A and Hb-F are both present in variable ratios in the same red cells. Moreover, under suitable conditions, all the red cells of sickle-cell anemic patients show the sickling phenomenon. Since relatively large amounts of Hb-F may be present in sickle-cell anemia, this prove sthat Hb-S and Hb-F are present together in the same red cells.

### 1. *The Occurrence of Hb-F in Adults*

Hb-F is present in some adults affected by hereditary or acquired hematological disorders. Hereditary conditions in which Hb-F is constantly found in adults include sickle-cell anemia, thalassemia major, and the anemias caused by a thalassemic gene in combination with a gene for an abnormal hemoglobin. Acquired pathological conditions in which Hb-F has occasionally been reported include pernicious anemia, hypoplastic anemia, leukemia, and other types of anemia (Singer *et al.*, 1951; Chernoff, 1953; White and Beaven, 1959).

The level of Hb-F in anemic patients varies over a wide range; the highest levels of Hb-F have been observed in thalassemia major, where Hb-F frequently approaches 90–95% (Rich, 1952). A slight amount of Hb-F has been reported to be present in pregnant women (Rucknagel and Chernoff, 1955) and in women with hydatid moles (Bromberg *et al.*, 1957), conditions in which high amounts of chorionic gonadotropin are secreted. These observations suggested that gonadotropin may stimulate Hb-F production (Rucknagel and Chernoff, 1955).

The appearance of Hb-F in hematological disorders has been considered to be the consequence of severe and chronic anemia or a com-

---

FIG. 13. The switch from Hb-F synthesis to Hb-A synthesis. ——— indicates the mean relative amount of $\gamma$ peptide chain present during the last months of fetal life and the first 10 months after birth; ------- indicates the mean relative amount of $\beta$ peptide chain; and — — — — the relative amount of $\alpha$ peptide chain during the same period. The $\alpha$ peptide chain is constantly 100% since this chain is common to Hb-F and Hb-A; • • • • • = percentages of Hb-F present at birth in relation to gestational age (data from Cook *et al.*, 1957). x x x x x = percentages of Hb-F present in children from 1 to 10 months of age (data from Chernoff and Singer, 1952).

pensatory phenomenon when the synthesis of Hb-A is impaired, as in thalassemia major (Rich, 1952).

Hb-F does not seem to be evenly distributed in the red cells of anemic patients. Bradley *et al.* (1961) have reported that Hb-F and Hb-S are unevenly distributed in the red cells of sickle-cell anemic patients. Similar results have been obtained by the Kleihauer *et al.* (1957) method of analysis by Thompson *et al.* (1961) and by Mitchener *et al.* (1961). Identical observations have been reported for the distribution of Hb-F in red cells of subjects affected by thalassemia major (Thompson *et al.*, 1961). Two distinct cell populations are observed in these patients by this method of analysis, one with high amounts of Hb-F and the other with rather small amounts of Hb-F (Thompson *et al.*, 1961). The Hb-F in the red cells of severe anemic patients seem thus to have a clonal distribution, with clones of red cells containing Hb-F and clones of red cells apparently of the normal adult type.

### C. HEREDITARY PERSISTENCE OF Hb-F

Edington and Lehmann (1955), and subsequently other investigators, have reported a condition in which high proportions of Hb-F are found in healthy Negro adults. This condition has been designated "hereditary persistence of Hb-F" by Jacob and Raper (1958), since Mendelian inheritance of this anomaly was shown by family studies. The gene responsible for the persistence of Hb-F during adult life has also been called "high F" gene. In carriers of this gene 20 to 30% of the hemoglobin is Hb-F; approximately the same amount of Hb-F is observed in individuals heterozygous for the high F gene and for the abnormal hemoglobin genes $\beta^{S}$ or $\beta^{C}$. It is quite remarkable that no Hb-A is present in these heterozygous individuals (Jacob and Raper, 1958; Went and MacIver, 1958).

Several marriages of heterozygotes for the high F and $\beta^{S}$ or $\beta^{C}$ genes with normal partners have been reported (Bradley *et al.*, 1961; Thompson *et al.*, 1961; Kraus *et al.*, 1961; MacIver *et al.*, 1961); children born from these marriages inherit either the high F gene or the abnormal hemoglobin gene.

The high F gene thus behaves as if allelic or closely linked to the $\beta$ locus. It has been suggested that the high F abnormality is the result of a mutation at the $\beta$ locus and that the production of Hb-F is a compensatory phenomenon (Bradley and Conley, 1960). The high F gene in this view is interpreted as a thalassemic-like gene, suppressing completely the production of $\beta$ peptide chains. However, the majority of the heterozygotes for a thalassemic gene do not show any increase of the Hb-F level over normal values. Individuals heterozygous for the high

F gene and a $\beta$ thalassemic gene have been reported (Kraus *et al.*, 1961). The high F gene seems to interact with the $\beta^{Th}$ gene; 70% Hb-F and 30% Hb-A are present in these individuals. Similar levels of Hb-F are normally found in thalassemia major; but the high F/thalassemia condition is of intermediate severity and quite different from thalassemia major as a pathological entity (Kraus *et al.*, 1961).

Individuals with the high F sickle-cell condition have levels of Hb-F comparable to some of the sickle-cell anemic patients (Bradley *et al.*, 1961); the ratio Hb-S/Hb-F is, however, quite variable in sickle-cell anemia. Sickle-cell anemia is a severe disease, while the high F/sickle-cell condition is almost asymptomatic (Edington and Lehmann, 1955; Jacob and Raper, 1958; Went and MacIver, 1958). Thompson *et al.* (1961) and Bradley *et al.* (1961) have reported that Hb-F is evenly distributed in the red cells of high F and of high F/sickle-cell individuals, in contrast with the uneven distribution of Hb-F in thalassemia major and sickle-cell anemia. This observation provides a clue to explain the absence of pathological manifestations in the high F/sickle cell condition, as compared to sickle-cell anemia. In sickle-cell anemia the cells with high Hb-S content sickle and lyse at normal physiological oxygen tension, while in the high F/sickle-cell condition the uniform distribution of Hb-F among the red cells prevents sickling at the same oxygen tension. Hb-F does not participate in the sickling process, since it does not form mixed aggregates with Hb-S (Allison, 1957).

The uniform distribution of Hb-F in red cells of high F individuals is of more profound significance, since it shows that the synthesis of high amounts of Hb-F is common to all the cells and is not characteristic of particular cellular clones, as it seems to be the case in severe anemias. The synthesis of Hb-F in the high F condition is the primary manifestation of the "high F" gene and does not seem to result indirectly from severe anemia (Thompson *et al.*, 1961).

Wheeler and Krevans (1961) have reported a child apparently homozygous for the high F gene; the child is in healthy condition and does not possess any detectable Hb-A or Hb-$A_2$. The absence of Hb-$A_2$ in this child constitutes the only evidence that the high F gene suppresses the synthesis of $\beta$ and $\delta$ chains at the same time. The Hb-F of this child appears to be identical by fingerprinting to authentic Hb-F from cord blood (C. Baglioni, unpublished).

The hereditary persistence of Hb-F seems to be a genetic entity completely different from the hereditary anemias in which high levels of Hb-F have been reported. It has been suggested that the high F mutation may involve a regulator or operator gene, rather than a structural gene (P. S. Gerald, unpublished; Ceppellini, 1961; Motulsky, 1962; Neel, 1961). The function of such a regulator gene is to turn off the $\gamma$ chain production and to turn on the $\beta$ and $\delta$ chain production. The fact that individuals heterozygous for the high F gene and for a $\beta^S$ or a $\beta^C$ gene form Hb-S or Hb-C in normal amount seems to indicate that the regulator gene exerts its control only on the structural genes in *cis,* located on the same chromosome (Neel, 1961). The suppression of $\delta$ chains production in the child homozygous for the high F gene (Wheeler and Krevans, 1961) is in agreement with the idea that the regulator gene extends its control over the $\beta$ and $\delta$ structural genes, which are known to be closely linked (Ceppellini, 1959b).

The mode of action of the high F gene is analogous in form with that of the operator gene of Jacob and Monod (1961). The operator gene has been defined by Jacob and Monod (1961) as a genetic element which controls the expression of several adjacent structural genes; the operator as defined in the *Lac* system of Jacob and Monod is the site of action of repressors produced by regulator genes. The interaction of regulators, operators, and structural genes results in the repressive genetic regulation of protein synthesis.

In Fig. 14 the high F mutation has been indicated as $o^-$ and the normal allele as $o^+$, by analogy with the nomenclature of Jacob and Monod (1961). The close linkage of the postulated $o$ gene with the $\beta$ and $\delta$ genes is in agreement with the structure of the coordinated unit of genetic expression defined by Jacob and Monod (1961) as *operon.* It must be realized that the existence of an operator among the hemoglobin genes is not substantiated by genetic evidence minimally comparable to that obtained by the fine genetic analysis of the *Lac* region of *E. coli.* It is extremely useful, however, to borrow Jacob and Monod's model of the *operon* to explain tentatively the chemical findings for the high F condition. This model provides the investigator with a challenging hypothesis, and may be apt to stimulate further experimentation and criticism.

In the operator gene model it is hard to understand the relationships between concomitant events: the suppression of $\beta$ (and $\delta$) chain production and the persistent synthesis of $\gamma$ chains. In order to explain these apparently linked events, one has to assume that the $\beta$-$\delta$ operator, which is controlled via repression by a regulator gene, in its turn controls the $\gamma$ operator. In this way the synthesis of $\gamma$ chains may be shut off at the same time that the synthesis of $\beta$ and $\delta$ chains is activated. However, only one $\gamma$ gene, presumably the one in *cis* if $\gamma$ is linked to $\beta$ and $\delta$, should fall under this control, since the normal operator in the heterozygotes does not suppress $\gamma$ chain production. Nothing is known about the position of the $\gamma$ locus, since $\gamma$ alleles have not been found. The operator gene hypothesis is still highly speculative in the case of the human hemoglobin genes and needs to be substantiated by more direct and critical evidence.

It cannot be excluded on genetic grounds that the high F condition results from an inactivation or a deletion of the two linked genes β and δ, however unlikely a recurrent genetic event of this type may be; the concomitant elevation of Hb-F still remains to be explained. It has been suggested (V. M. Ingram, unpublished) that the difference between thalassemia and high F condition may reside in the production of a defective RNA-template in the former and in no production at all of RNA-templates in the latter, if the β and δ genes have been deleted. If messenger-RNA molecules compete for the ribosomal sites of protein synthesis, in thalassemia the ribosomes would be occupied by the defective messengers, while in the high F condition the ribosomes may be free to accept γ chain RNA-templates.

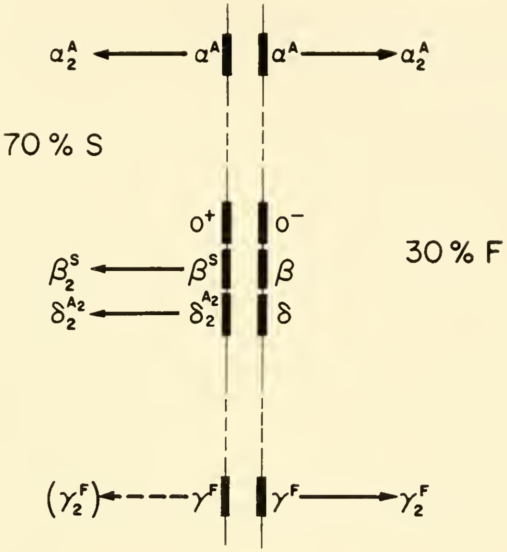


FIG. 14. Schematic representation of the postulated genotype of individuals heterozygous for the "hereditary persistence of Hb-F" and for the $\beta^S$ gene. The normal "operator" or "regulator" gene (Neel, 1961) is indicated as $o^+$; the abnormal allele responsible for the persistence of Hb-F is indicated as $o^-$. The arrows illustrate the dimers synthesized under the control of the hemoglobin genes and the hemoglobins produced by random combination of the dimers (capital letters). The β and δ genes in *cis* with the $o^-$ mutation are supposedly inactive (see Section V,C); it is not known whether γ chains are synthesized under the control of the γ locus in *cis* with the $o^-$ mutation or whether both the γ loci are active. The relative levels of Hb-S and Hb-F observed in these individuals are also shown.

### D. MODELS OF RED CELL DIFFERENTIATION AND SYNTHESIS OF Hb-F

Red cells are formed through a process of differentiation from nucleated precursors. A continuous series of morphologically well-defined intermediate stages of differentiation between undifferentiated precursors (stem cells) and red cells has long been recognized (see Wintrobe, 1962). Several models of stem cell differentiation have been proposed recently to explain the kinetics of red cell production and its modifications in pathological conditions (Alpen and Cranmore, 1959; Stohlman, 1959; Lajtha and Oliver, 1960). Every model of red cell proliferation indicates the necessity for the perpetuation of the primitive stem cells in the absence of differentiation. The stem cells appear to be morphologically and biochemically undifferentiated; they act as a reservoir of potential blood cells. Only a small proportion of stem cells per unit of time undergoes differentiation normally, while replication of undifferentiated stem cells provides the somatic continuity of the cell line.

Barnes *et al.* (1959) have transplanted $10^6$ nucleated cells from the spleen of an infant $T6^+$ mouse to an irradiated CBA mouse; from the bone marrow of the latter the same number of cells were serially transplanted into irradiated CBA mice, succeeding each time in populating the bone marrow of the CBA mice with cytologically distinguishable $T6^+$ cells. This experiment provides direct experimental evidence of the unlimited proliferative ability of the stem cells.

It has been proposed by Alpen and Cranmore (1959) that the normal supply of red cells is, in large part, provided by the division of already differentiated erythroid precursors, called erythroblasts or pronormoblasts. Since the synthesis and the accumulation of hemoglobin initiates soon after the stem cells differentiate into erythroid precursors, these cells can be specifically labeled by means of the radioactive isotope $Fe^{59}$. The erythroblast is the earliest cell type that incorporates measurable amounts of $Fe^{59}$; the label incorporated can be estimated by autoradiographic techniques in which the grains produced by single cells are counted; division of the cells causes halving of the grain count. Alpen and Cranmore (1959) have observed that some of the erythroblasts divide without differentiating further and suggested that the erythroblasts may divide into a nearly equal number of cells which will differentiate further, and of cells which do not differentiate but provide the continuity of the cell line.

In animals made anemic by bleeding, a large number of stem cells differentiate rapidly into erythroblasts (Lajtha and Oliver, 1960) to compensate for the increased demand in red cell output. Whereas 24 hours after $Fe^{59}$ injection 100% of the erythroblasts are labeled in normal animals, only 10% are found to be labeled in bled animals (Alpen and Cranmore, 1959); the average generation time of the erythroblast does not appear to be different in bled or control animals. The erythroblasts seem to be a partially self-maintaining cell population, supplemented by the differentiation of the stem cells. This implies that the erythroblasts are able to make a choice between dividing only or dividing and differentiating.

A definite correlation exists between higher erythropoietic activity in pathological conditions and the appearance of Hb-F in adults. Although we are far from a full understanding of the process of red cell differentiation, some of the above-mentioned facts suggest a possible explanation for the switch from Hb-F to Hb-A production at the end of fetal life and for the synthesis of Hb-F in anemic individuals.

It may be assumed that the hemoglobin type synthesized by a red cell precursor is determined by the genetic information possessed by that cell at the moment of initiating hemoglobin synthesis and during the course of hemoglobin synthesis. By analogy with protein synthesis in bacteria, we can speculate that the hemoglobin genes are repressed in stem cells and that when these cells differentiate into erythroblasts, specific hemoglobin genes become derepressed or activated by genetically determined mechanisms. The factors determining derepression or activation of $\beta$ rather than $\gamma$ genes are unknown. It has been suggested that the change in oxygen tension from the uterine environment to the extrauterine environment after birth may be the activating mechanism for the synthesis of $\beta$ chains (Allen and Jandl, 1960; Thomas *et al.,* 1960) or that gonadotropic hormones may play a similar role (Rucknagel and Chernoff, 1955). None of these proposed mechanisms accounts satisfactorily, in the author's opinion, for the synthesis of Hb-F in pathological conditions of severe anemia during adult life.

By analogy with the observations in bled animals, we can speculate that in anemic subjects the stem cells differentiate into red cells after a limited number of cell divisions, while in normal erythropoiesis a considerable number of cell divisions occur at the erythroblast stage before maturation of the red cells. It has been reported that stem cells may sometimes differentiate into red cells without intervening division (Suit, 1957). Similarly, in the fetus the red cells are released in circulation after a number of divisions of the erythroid precursors, limited by the age of the fetus itself and possibly by the accelerated fetal erythropoietic activity. A correlation may thus exist either between the number of divisions that the progenitors of a red cell have undergone in the absence of morphological differentiation or the time spent in the bone marrow by these cells as erythroblasts and the hemoglobin type con-

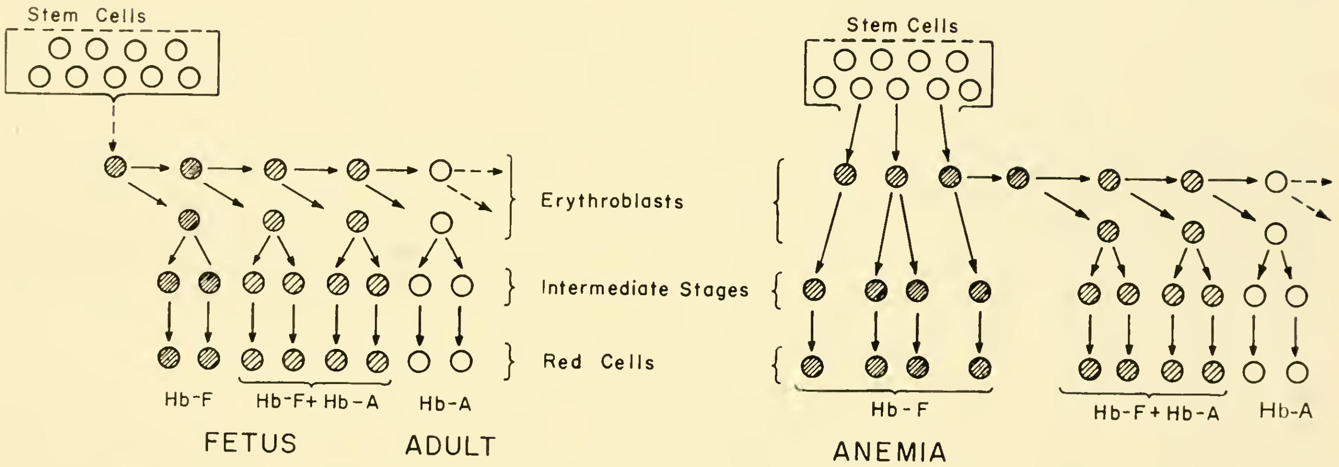

Stem Cells
Erythroblasts
Intermediate Stages
Red Cells
Hb-F
Hb-F+ Hb-A
Hb-A
FETUS
ADULT
Stem Cells
Hb - F
Hb-F + Hb-A
Hb-A
ANEMIA

FIG. 15. Model of red cell differentiation, showing the switch from Hb-F to Hb-A production at the end of fetal life (left) and the mechanism by which Hb-F is produced in severe anemia (right). (See Section V,D for further explanations and for references.) The intensity of the shading illustrates the relative amount of Hb-F synthesized by red cell precursors at different stages of differentiation or contained in red cells.

tained by the daughter red cell. Red cells produced after few cell divisions from stem cells may possibly contain more Hb-F than red cells produced after several divisions of the erythroblasts. In this view the information directed to the hemoglobin chain synthesis changes with the aging and/or the replication of the erythroid precursors (Fig. 15).

Sublethal irradiation of monkeys (Tuttle *et al.*, 1961) or of humans for therapeutic purposes (R. Ceppellini, personal communication), which causes rapid differentiation of stem cells (Lajtha and Oliver, 1960), has been reported to be followed by the appearance of Hb-F in the blood of the irradiated subjects.

The present hypothesis regarding the relationship between red cell differentiation and the switch from Hb-F to Hb-A production accounts formally for most of the observations made in recent years: the lack of effect of birth on the switch from Hb-F to Hb-A production (Cook *et al.*, 1957) and the appearance of Hb-F in conditions of abnormal erythropoiesis. It fails, however, to provide an explanation of the mechanisms involved in the control of the hemoglobin genes. This seems to be a point of more general interest since it deals with the control of differentiation, the challenging problem of today's biology.

The Hb-F/Hb-A system may be useful in the study of mechanisms of differentiation, because of the well-defined chemical nature of the proteins involved. In the switch from Hb-F to Hb-A production, only the synthesis of one type of peptide chain, the $\gamma$ chain, seems to be discontinued, while the synthesis of a new type of peptide chain, the $\beta$ chain, is initiated. Only two genes, the $\beta$ and the $\gamma$ genes, are apparently involved in switching their activity; this fact may provide a unique opportunity for the study of the differentiation of a simplified genetic and biochemical system.

## Acknowledgments

The author wishes to thank Dr. Vernon M. Ingram for reading this manuscript, for suggesting several improvements and corrections, and for discussion and advice. Thanks are due to all who have made available unpublished data and who have contributed to this chapter with personal communications. The author is grateful to the Medical Foundation, Inc., Boston, and to the National Science Foundation, Washington, who have supported his work with grants.

## References

Ager, J. A. M., and Lehmann, H. (1958). *Brit. Med. J.* **I**, 929.
Aksoy, M., and Lehmann, H. (1957). *Nature* **179**, 1248.
Allen, D. W., and Jandl, J. H. (1960). *J. Clin. Invest.* **39**, 1107.
Allen, D. W., Schroeder, W. A., and Balog, J. (1958). *J. Am. Chem. Soc.* **80**, 1628.
Allison, A. C. (1955). *Science* **122**, 640.
Allison, A. C. (1957). *Biochem. J.* **65**, 212.

Alpen, E. L., and Cranmore, D. (1959). *In* "The Kinetics of Cellular Proliferation" (F. Stohlman, Jr., ed.), p. 290. Grune & Stratton, New York and London.

Atwater, J., Schwartz, I. R., and Tocantins, L. M. (1960a). *Blood* **15**, 901.

Atwater, J., Schwartz, I. R., Erslev, A. J., Montgomery, T. L., and Tocantins, L. M. (1960b). *New Engl. J. Med.* 263, 1215.

Baglioni, C. (1961). *Biochim. et Biophys. Acta* **48**, 392.

Baglioni, C. (1962a). *J. Biol. Chem.* **237**, 69.

Baglioni, C. (1962b). *Biochim. et Biophys. Acta* **59**, 437.

Baglioni, C., and Ingram, V. M. (1961). *Biochim. et Biophys. Acta* **48**, 253.

Baglioni, C., and Lehmann, H. (1962). *Nature,* in press.

Baglioni, C., Ingram, V. M., and Sullivan, E. (1961). *Nature* **189**, 467.

Bailey, I. S., and Prankerd, T. A. J. (1958). *Brit. J. Haematol.* **4**, 150.

Barnes, D. W. H., Ford, C. E., and Loutit, J. F. (1959). *Sang* **30**, 762.

Benesch, R. E., Ranney, H. M., Benesch, R., and Smith, G. M. (1961). *J. Biol. Chem.* **236**, 2926.

Benzer, S., Ingram, V. M., and Lehmann, H. (1958). *Nature* **182**, 852.

Betke, K. (1958). *Schweiz. med. Wochschr.* **88**, 1005.

Boulard, C., Cosset, A., Destaing, F., Duzer, A., Jonxis, J. H. P., Muller, C. J., and Portier, A. (1961). *Blood* **18**, 750.

Bradley, T. B., and Conley, C. L. (1960). *Trans. Assoc. Am. Physicians* **73**, 72.

Bradley, T. B., Jr., Brawner, J. N., III, and Conley, C. L. (1961). *Bull. Johns Hopkins Hosp.* **108**, 242.

Braunitzer, G., Rudloff, V., Hilse, K., Liebold, B., and Müller, R. (1960). *Z. physiol. Chem. Hoppe-Seyler's* **320**, 283.

Braunitzer, G., Hillschmann, N., Rudloff, V., Hilse, K., Liebold, B., and Müller, R. (1961a). *Nature* **190**, 480.

Braunitzer, G., Gehring-Müller, R., Hillschmann, N., Hilse, K., Hobam, G., Rudloff, V., and Wittman-Liebold, B. (1961b). *Z. physiol. Chem. Hoppe-Seyler's* **325**, 283.

Bridges, C. B. (1936). *Science* **83**, 210.

Bromberg, Y. M., Yehuda, M., Salzberger, M., and Abrahamov, A. (1957). *Blood* **12**, 1122.

Cabannes, R., and Portier, A. (1959). *In* "Abnormal Haemoglobins" (J. H. P. Jonxis and J. F. Delafresnaye, eds.), p. 51. Blackwell, Oxford.

Carcassi, V., Ceppellini, R., and Siniscalco, M. (1957). *Haematologica Pavia* **42**, 1635.

Ceppellini, R. (1956). *Sangue* **1**, 1.

Ceppellini, R. (1959a). *Acta Genet. Med. et Gemellol.* **8**, 47.

Ceppellini, R. (1959b). *Ciba Foundation Symposium, Biochemistry of Human Genetics,* p. 133.

Ceppellini, R. (1961). *2nd Intern. Conf. Human Genet. Rome.*

Charlwood, P. A., Gratser, W. B., and Beaven, G. H. (1960). *Biochim. et Biophys. Acta* **40**, 191.

Chernoff, A. I. (1953). *Blood* **8**, 413.

Chernoff, A. I. (1956). *New Engl. J. Med.* 253, 322 (1955).

Chernoff, A. I., and Liu, J. C. (1961). *Blood* **17**, 54.

Chernoff, A. I., and Singer, R. (1952). *Pediatrics* **9**, 469.

Clegg, M. D., and Schroeder, W. A. (1959). *J. Am. Chem. Soc.* **81**, 6065.

Cook, C. D., Brodie, H. R., and Allen, D. W. (1957). *Pediatrics* **20**, 272.

Crick, F. H. C., Barnett, L., Brenner, S., and Watts-Tobin, R. J. (1961). *Nature* **192**, 1227.

Dintzis, H. M. (1961). *Proc. Natl. Acad. Sci.* (*U.S.*) **47**, 247.

Drabkin, D. L., and Austin, J. H. (1935). *J. Biol. Chem.* **112**, 89.

Drescher, H., and Künzer, W. (1954). *Klin. Wochschr.* 32, 92.

Dreyer, W. J. (1960). *Brookhaven Symposia in Biol.* 13, 243.

Edington, G. M., and Lehmann, H. (1955). *Brit. Med. J.* **II**, 1328.

Fessas, P. (1959). *In* "Abnormal Haemoglobins" (J. H. P. Jonxis and J. F. Delafresnaye, eds.), p. 153. Blackwell, Oxford.

Fessas, P., and Papaspyrou, A. (1957). *Science* **126**, 1119.

Fessas, P., Karaklis, A., and Gnafakis, N. (1961). *Acta Haematol.* **25**, 62.

Fessas, P., Stamatoyannopoulos, G., and Karaklis, A. (1962). *Blood* **19**, 1.

Field, E. O., and O'Brien, J. A. P. (1955). *Biochem. J.* **60**, 656.

Freese, E. (1958). *Cold Spring Harbor Symp. Quant. Biol.* **23**, 13.

Gammack, D. B., Huehns, E. R., Shooter, E. M., and Gerald, P. S. (1960). *J. Mol. Biol.* **2**, 372.

Gammack, D. B., Huehns, E. R., Lehmann, H., and Shooter, E. M. (1961). *Acta Genet. et Statist. Med.* **11**, 1.

Garen, A., Levinthal, C., and Rothman, F. (1961). *J. chim. phys.* **58**, 1068.

Gerald, P. S., and Diamond, L. K. (1958). *Blood* **13**, 835.

Gerald, P. S., and Efron, M. L. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 1758.

Gerald, P. S., and Ingram, V. M. (1961). *J. Biol. Chem.* **236**, 2155.

Gerald, P. S., Efron, M., and Diamond, L. K. (1961). *Am. J. Diseases Children* **102**, 514.

Goldstein, J., Guidotti, G., Konigsberg, W., and Hill, R. J. (1961). *J. Biol. Chem.* **236**, PC77.

Gouttas, A., Fessas, P., Tsevremis, H., and Xefteri, E. (1955). *Sang* **26**, 911.

Gratzer, W. B., and Allison, A. C. (1960). *Biol. Rev.* **35**, 459.

Gratzer, W. B., and Beaven, G. H. (1961). *J. Chromatog.* **5**, 315.

Halbrecht, J., and Klibansky, C. (1956). *Nature* **178**, 794.

Hasserodt, V., and Vinograd, J. (1959). *Proc. Natl. Acad. Sci. U.S.* **45**, 12.

Helinski, D. R., and Yanofsky, C. (1962). *Proc. Natl. Acad. Sci. U.S.* 48, 173.

Hendrickse, R. G., Boyo, A. E., Fitzgerald, P. A., and Ransome, S. (1960). *Brit. Med. J.*, **I** 611.

Hill, R. J., and Craig, L. C. (1959). *J. Am. Chem. Soc.* **81**, 2272.

Hill, R. L., and Schwartz, H. C. (1959). *Nature* **184**, 642.

Hill, R. J., and Konigsberg, W. (1960). *J. Biol. Chem.* **235**, PC21.

Hill, R. J., and Konigsberg, W. (1961). *J. Biol. Chem.* **236**, PC7.

Hill, R. L., Swenson, R. T., and Schwartz, H. C. (1960). *J. Biol. Chem.* **235**, 3182.

Hitzig, W. H., Frick, P. G., Betke, K., and Huisman, T. H. J. (1960). *Helv. Paediat. Acta* **15**, 499.

Hoch, H. (1950). *Biochem. J.* **46**, 199.

Hörlein, H., and Weber, G. (1948). *Deut. med. Wochschr.* 39, 476.

Horton, B., Payne, R. A., Bridges, M. T., and Huisman, T. H. J. (1961). *Clin. Chim. Acta* **6**, 246.

Huehns, E. R., and Shooter, E. M. (1961). *Nature* **189**, 918.

Huehns, E. R., Flynn, F. V., Butler, E. A., and Beaven, G. H. (1961a). *Nature* **189**, 496.

Huehns, E. R., Shooter, E. M., Dance, N., Beaven, G. H., and Shooter, K. V. (1961b). *Nature* **192**, 1057.

Huisman, T. H. J. (1960). *Nature* **188**, 589.

Huisman, T. H. J. (1961). *Biochim. et Biophys. Acta* **46**, 384.

Huisman, T. H. J., and Meyering, C. A. (1960). *Clin. Chim. Acta* **5**, 103.

Huisman, T. H. J., Punt, K., and Schaad, J. D. G. (1961). *Blood* **17**, 747.
Hunt, J. A. (1959). *Nature* **183**, 1373.
Hunt, J. A., and Ingram, V. M. (1958a). *Biochim. et Biophys. Acta* **28**, 546.
Hunt, J. A., and Ingram, V. M. (1958b). *Nature* **181**, 1062.
Hunt, J. A., and Ingram, V. M. (1959). *Ciba Foundation Symposium, Biochemistry of Human Genetics,* p. 114.
Hunt, J. A., and Ingram, V. M. (1960). *Biochim. et Biophys. Acta* **42**, 409.
Hunt, J. A., and Lehmann, H. (1959). *Nature* **184**, 872.
Ingram, V. M. (1956). *Nature* **178**, 792.
Ingram, V. M. (1958). *Biochim. et Biophys. Acta* **28**, 539.
Ingram, V. M. (1959a). *Nature* **183**, 1795.
Ingram, V. M. (1959b). *Biochim. et Biophys. Acta* **36**, 402.
Ingram, V. M. (1961a). *Nature* **189**, 704.
Ingram, V. M. (1961b). *In* "Hemoglobin and Its Abnormalities." C. C Thomas, Springfield, Illinois.
Ingram, V. M., and Stretton, A. O. W. (1959a). *Federation Proc.* **19**, 343.
Ingram, V. M., and Stretton, A. O. W. (1959b). *Nature* **184**, 1903.
Ingram, V. M., and Stretton, A. O. W. (1961). *Nature* **190**, 1079.
Ingram, V. M., and Stretton, A. O. W. (1962). *Biochim. et Biophys. Acta* **62**, 456; **63**, 20.
Itano, H. A. (1951). *Proc. Natl. Acad. Sci. U.S.* **37**, 775.
Itano, H. A. (1953). *Am. J. Human Genet.* **5**, 34.
Itano, H. A. (1957). *Advances in Protein Chem.* **12**, 215.
Itano, H. A., and Neel, J. V. (1950). *Proc. Natl. Acad. Sci. U.S.* **36**, 613.
Itano, H. A., and Robinson, E. A. (1959). *Nature* **183**, 1799.
Itano, H. A., and Robinson, E. A. (1960). *Proc. Natl. Acad. Sci.* **46**, 1492.
Itano, H. A., and Singer, S. J. (1958). *Proc. Natl. Acad. Sci. U.S.* **44**, 522.
Itano, H. A., Bergren, W. R., and Sturgeon, P. (1954). *J. Am. Chem. Soc.* **76**, 2278.
Itano, H. A., Singer, S. J., and Robinson, E. (1959). *Ciba Foundation Symposium, Biochemistry of Human Genetics,* p. 96.
Jacob, F., Monod, J. (1961). *J. Mol. Biol.* **3**, 318.
Jacob, G. F., and Raper, A. B. (1958). *Brit. J. Haematol.* **4**, 138.
Jones, R. T., Schroeder, W. A., Balog, J. E., and Vinograd, J. R. (1959a). *J. Am. Chem Soc.* **81**, 3161.
Jones, R. T., Schroeder, W. A., and Vinograd, J. R. (1959b). *J. Am. Chem. Soc.* **81**, 4749.
Jonxis, J. H. P. (1959). *In* "Abnormal Haemoglobins" (J. H. P. Jonxis and J. F. Delafresnaye, eds.), p. 114. Blackwell, Oxford, England.
Jonxis, J. H. P., and Visser, H. K. A. (1956). *Am. J. Diseases Children* **92**, 588.
Kendrew, G. C., Watson, H. C., Strandberg, B. E., Dickerson, R. E., Phillips, D. C., and Shore, V. C. (1961). *Nature* **190**, 666.
Kleihauer, E., Braun, H., and Betke, K. (1957). *Klin. Wochschr.* **35**, 637.
Konigsberg, W., Guidotti, G., and Hill, R. J. (1961). *J. Biol. Chem.* **236**, PC55.
Kraus, A. P., Koch, B., and Burckett, L. (1961). *Brit. Med. J.* 1434.
Kunkel, H. G., and Bearn, A. G. (1957). *Federation Proc.* **16**, 760.
Kunkel, H. G., Ceppellini, R., Müller-Eberhardt, V., and Wolf, J. (1957). *J. Clin. Invest.* **36**, 1615.
Kunkel, H. G., and Wallenius, G. (1955). *Science* **122**, 288.
Lajtha, L. G., and Oliver, R. (1960). *Ciba Foundation Symposium, Haemopoiesis: Cell Production and Its Regulation* p. 289.
Lehmann, H. (1960). *In* "The Metabolic Basis of Inherited Disease" (J. B. Stan-

bury, J. B. Wyngaarden, and D. S. Fredrickson, eds.), p. 1086. McGraw-Hill, New York.

Lengyel, P., Speyer, J. F., and Ochoa, S. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 1936.

Liebold, B., and Braunitzer, G. (1959). *Z. physiol. Chem. Hoppe-Seyler's* **315**, 271.

Lie-Injo, L. E. (1961). *Nature* **191**, 1314.

Lie-Injo, L. E., and Lie, H. G. (1961). *Acta Haematol.* **25**, 192.

MacIver, J. E., Went, L. N., and Irvine, R. A. (1961). *Brit. J. Haematol.* **7**, 373.

Matsuda, G., Schroeder, W. A., Jones, R. T., and Weliky, N. (1960). *Blood* **16**, 984.

Matthaei, J. H., Jones, O. W., Martin, R. G., and Nirenberg, M. W. (1962). *Proc. Natl. Acad. Sci. U.S.* **48**, 666.

Maximow, A. A., and Bloom, W. (1957). "Textbook of Histology," 7th ed. Saunders, Philadelphia.

Minnich, V., Cordonnier, J. K., Williams, W. J., and Moore, C. V. (1962). *Blood* **19**, 137.

Mitchener, J. W., Thompson, R. B., and Huisman, T. H. J. (1961). *Lancet* **i**, 1169, 1349.

Morrison, M., and Cook, J. L. (1955). *Science* **122**, 920.

Motulsky, A. G. (1956). *Nature* **178**, 1055.

Motulsky, A. G. (1962). *Nature* **194**, 607.

Muller, C. J. (1961). A comparative study of the structure of mammalian and avian haemoglobins. Doctoral Thesis. Groningen, Holland.

Muller, C. J., and Jonxis, J. H. P. (1960). *Nature* **188**, 949.

Muller, C. J., and Kingma, S. (1961). *Biochim. et Biophys. Acta* **50**, 595.

Murayama, M., and Ingram, V. M. (1959). *Nature* **183**, 1798.

Neeb, H., Beiboer, J. L., Jonxis, J. H. P., Kaars Sijpesteijn, J. A., and Muller, C. J. (1961). *Trop. geogr. Med.* **13**, 207.

Neel, J. V. (1949). *Science* **110**, 64.

Neel, J. V. (1956). *Ann. Human Genet.* **21**, 1.

Neel, J. V. (1961). *Blood* **18**, 769.

Nirenberg, M. W., and Matthaei, J. H. (1962). *Proc. Natl. Acad. Sci. U.S.* **48**, 1588.

Pauling, L., Itano, H. A., Singer, S. J., and Wells, I. C. (1949). *Science* **110**, 543.

Perutz, M. F., Rossman, M. G., Cullis, A. F., Muirhead, H., Will, G., and North, A. C. T. (1960). *Nature* **185**, 416.

Peterson, E. A., and Sober, H. A. (1956). *J. Am. Chem. Soc.* **78**, 751.

Poulik, M. D. (1957). *Nature* **180**, 1477.

Prins, H. K., and Huisman, T. H. J. (1956). *Nature* **177**, 840.

Ranney, H. M. (1954). *J. Clin. Invest.* **33**, 1634.

Raper, A. B., Gammack, D. B., Huehns, E. R., and Shooter, E. M. (1960). *Brit. Med. J.* **II**, 1257.

Rhinesmith, H. S., Schroeder, W. A., and Pauling, L. (1957a). *J. Am. Chem. Soc.* **79**, 609.

Rhinesmith, H. S., Schroeder, W. A., and Pauling, L. (1957b). *J. Am. Chem. Soc.* **79**, 4682.

Rhinesmith, H. S., Schroeder, W. A., and Martin, N. J. (1958). *J. Am. Chem. Soc.* **80**, 3358.

Rich, A. (1952). *Proc. Natl. Acad. Sci. U.S.* **38**, 187.

Rigas, D. A., Koler, R. D., and Osgood, E. E. (1955). *Science* **121**, 372.

Robinson, A. R., Zuelzer, W. W., Neel, J. V., Livingstone, F. B., and Miller, M. L. (1956). *Blood* **11**, 902.

Rucknagel, D. L., and Chernoff, A. I. (1955). *Blood* **10**, 1092.
Rucknagel, D. L., and Neel, J. V. (1961). *In* "Progress in Medical Genetics" (A. G. Steinberg, ed.), Vol. 1, p. 158. Grune & Stratton, New York and London.
Schnek, A. G., and Schroeder, W. A. (1961). *J. Am. Chem. Soc.* **83**, 1472.
Schroeder, W. A., and Matsuda, G. (1958). *J. Am. Chem. Soc.* **80**, 1521.
Schroeder, W. A., Jones, R. T., Shelton, J. R., Shelton, J. B., Cormick, J., and McCalla, K. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 811.
Schroeder, W. A., Shelton, J. R., Balog-Shelton, J., and Cormick, J. (1962). *Proc. Natl. Acad. Sci. U.S.* **48**, 284.
Scott, E. M., and Griffith, J. V. (1959). *Biochim. et Biophys. Acta* **34**, 584.
Shooter, E. M., Skinner, E. R., Garlick, J. P., and Barnicot, N. A. (1960). *Brit. J. Haematol.* **6**, 140.
Silvestroni, E., and Bianco, I. (1952). *Blood* **7**, 429.
Silvestroni, E., and Bianco, I. (1958). *Policlinico (Rome), Sez. prat.* **65**, 203.
Silvestroni, E., Bianco, I., Muzzolini, M., Modiano, G., and Vallisneri, E. (1957). *Progr. med.* **13**, 705.
Singer, K., Chernoff, A. I., and Singer, L. (1951). *Blood* **6**, 413, 429.
Singer, S. J., and Itano, H. A. (1959). *Proc. Natl. Acad. Sci. U.S.* **45**, 174.
Smith, E. L. (1962a). *Proc. Natl. Acad. Sci. U.S.* **48**, 677.
Smith, E. L. (1962b). *Proc. Natl. Acad. Sci. U.S.* **48**, 859.
Smith, E. W., and Torbert, J. V. (1958). *Bull. John Hopkins Hosp.* **101**, 38.
Smithies, O. (1959). *Advances in Protein Chemistry* **41**, 65.
Smithies, O., Dixon, G. H., and Connel, G. E. (1961). *2nd Intern. Conf. Human Genet. Rome.*
Speyer, J. F., Lengyel, P., Basilio, C., and Ochoa, S. (1962). *Proc. Natl. Acad. Sci. U.S.* **48**, 441.
Stohlman, F., Jr. (1959). *In* "The Kinetics of Cellular Proliferation" (F. Stohlman, Jr., ed.), p. 318. Grune & Stratton, New York and London.
Suit, H. D. (1957). *J. Clin. Pathol.* **10**, 267.
Swenson, R. T., Hill, R. L., Lehmann, H., and Jim, R. T. S. (1962). *J. Biol. Chem.* **237**, 1517.
Thomas, E. D., Lochte, H. L., Jr., Greenough, W. B., III, and Wales, M. (1960). *Nature* **185**, 396.
Thompson, R. B., Mitchener, J. W., and Huisman, T. H. J. (1961). *Blood* **18**, 267.
Thorup, O. A., Itano, H. A., Wheby, M., and Leavell, B. S. (1956). *Science* **123**, 889.
Tuttle, A. H., Newsome, F. E., Jackson, C. H., and Overman, R. R. (1961). *Science* **133**, 578.
Valentine, W. N., and Neel, J. V. (1944). *A.M.A. Arch. Internal Med.* **74**, 185.
Vella, F. (1959). *Nature* **184**, 272.
Vinograd, J. R., Hutchinson, W. D., and Schroeder, W. A. (1959). *J. Am. Chem. Soc.* **81**, 3168.
Walker, J., and Turnbull, E. P. N. (1955). *Arch. Disease Childhood* **30**, 11, 114.
Watson, H. C., and Kendrew, F. R. S. (1961). *Nature* **190**, 670.
Weatherall, D. J., and Baglioni, C. (1962). *Blood*, in press.
Weatherall, D. J., and Boyer, S. H. (1961). *Trans. Assoc. of Am. Physicians* **74**, 89.
Weatherall, D. J., and Boyer, S. H. (1962). *Bull. Johns Hopkins Hosp.* **110**, 8.
Wheeler, J. T., and Krevans, J. R. (1961). *Clin. Research* **9**, 168.
Went, L. N., and MacIver, J. E. (1958). *Blood* **13**, 559.
White, J. C., and Beaven, G. H. (1959). *Brit. Med. Bull.* **15**, 33.

Wilson, S., and Smith, D. B. (1959). *Canad. J. Biochem. and Physiol.* **37**, 405.

Wintrobe, M. M. (1962). *In* "Clinical Hematology," p. 38. Lea & Febiger, Philadelphia.

Zuelzer, W. W., and Kaplan, E. (1954). *Blood* **9**, 1047.

Zuelzer, W. W., Neel, J. V., and Robinson, A. R. (1956). *Progr. in Hematol.* **1**, 91.

## *Chapter X*

# Contributions from TMV Studies to the Problem of Genetic Information Transfer and Coding

A. TSUGITA[1] AND H. FRAENKEL-CONRAT

*Virus Laboratory,*
*University of California,*
*Berkeley, California*

I. Introduction . . . 477
II. TMV-Protein . . . 479
  A. Rod Structure and Chain Configuration . . . 479
  B. Amino Acid Sequence . . . 484
III. TMV-RNA . . . 487
IV. Viral Infectivity . . . 489
  A. Chemical Basis of Infectivity . . . 489
  B. Mechanism of Infection . . . 490
  C. Cell-Free Biosynthesis of TMV-Protein . . . 490
V. Natural Strains of TMV . . . 491
  A. Symptomatology of TMV and Its Strains . . . 491
  B. Composition of Natural Strains . . . 493
  C. Mixed Reconstitution . . . 495
  D. Structural Differences in the RNA of Different Strains . . . 496
VI. Chemically Produced Strains . . . 497
  A. Modification of the RNA . . . 497
  B. Protein Composition of Chemically Produced Mutants . . . 499
  C. Virulence and Viability of Strains . . . 508
VII. The Coding Problem . . . 509
  References . . . 517

## I. INTRODUCTION

The first description of a virus as a new kind of biological entity we owe to Iwanowski (1892), and particularly to Beijerinck (1899), who demonstrated that the tobacco mosaic disease was due to a replicating substance which was able to pass through bacterial filters. In the following 40 years many plant and animal diseases were attributed to this novel class of agents, termed the filtrable viruses. That bacteria could also harbor viruses was shown by Twort and by d'Hérelle about 1915. However, until about 1930 few chemists or biochemists concerned themselves with these biological agents which apparently were difficult to separate from cellular material.

---

[1] Present address: Research Institute for Protein, Osaka University, Osaka, Japan.

This situation changed when, in 1935, Stanley announced that he had isolated the tobacco mosaic virus in pure and crystalline form. The term virus then acquired a distinctive meaning in physicochemical terms. Soon several other plant viruses were isolated and crystallized, and all were proven to consist of RNA and protein (Bawden and Pirie, 1937; Stanley, 1938). Greater difficulties were encountered in the isolation of animal viruses, but several of them have also been obtained in crystalline form (Schaffer and Schwerdt, 1955). These, as well as the bacterial viruses, proved to contain either RNA or DNA. Through these developments viruses became the subject of intensive physical and chemical, as well as biological, research. With the application of new techniques for the study of macromolecules, such as electron microscopy, X-ray or light scattering, and ultracentrifugation, the structure of many viruses came to be well understood.

The general pattern emerged that most of the simple viruses, consisting only of RNA and protein, were either of approximately spherical dimensions, or rod-shaped. In all viruses, the nucleic acid was covered by the protein. The amount of protein per particle varied greatly, but the amount of RNA was approximately constant: 1–2 million "molecular weight" per virus particle ranging from about 5 million for the turnip yellows mosaic virus to 100 million for the influenza virus (Frisch-Niggemeyer, 1956).

In recent years it has become evident and generally accepted that in the case of many of the simpler viruses the RNA is by itself infectious even though of low efficiency (Fraenkel-Conrat, 1956; Gierer and Schramm, 1956). Furthermore, it has been clearly established that the RNA carries the entire genetic information required for virus production. Indications have been obtained that intracellular release of the nucleic acid is one of the first events upon virus infection with the RNA viruses, as was so well demonstrated to be the case with the bacteriophages (Hershey and Chase, 1952). Thus, virus particles are now generally regarded as representing relatively small-molecular genetic nucleic acids wrapped for transport in specifically designed and genetically determined protein overcoats of greater or lesser functional sophistication. In the simple RNA viruses, the role of the protein may be solely structural, but even here evolution has produced some surprising results. Thus, the complete protein coat may represent not only an adequate protective covering for the RNA, but may also give the particle properties of thermal stability and enzyme resistance which are remarkable, considering its predominant protein nature.

Tobacco mosaic virus (TMV) was the virus first to be discovered, first to be isolated, and first to be degraded into protein and infective RNA, and to be reconstituted from these components. Because so much more is known about this than most other viruses, this article will be largely devoted to TMV. We shall first discuss the principle features of the structure and function of the tobacco mosaic virus, and we shall then attempt to show in what manner the study of this virus has contributed and may continue to contribute to our understanding of the molecular aspects of genetics. More detailed discussions of various aspects of this field can be found in recent reviews (Fraenkel-Conrat and Ramachandran, 1959; Schuster, 1960; Klug and Caspar, 1960; Gierer, 1960).

## II. TMV-Protein

### A. ROD STRUCTURE AND CHAIN CONFIGURATION

Tobacco mosaic virus (TMV) is a rod-shaped particle (radius 150–180 Å, length 3000 Å) composed of 5% ribonucleic acid (RNA) and 95% protein and having a particle weight of 40 million. The model presented in Fig. 1 is largely based on data obtained through a thorough X-ray diffraction analysis (Franklin and Klug, 1955, 1956; Franklin *et al.*, 1957; Klug and Caspar, 1960), particularly of a methyl mercury derivative of the virus (Franklin and Holmes, 1958). As shown in the figure, the rod is believed to be hollow along its axis (40 Å diameter) and surrounded by protein units in a gently pitched helical array (49 units in 3 turns, pitch 23 Å). The nucleic acid is threaded through about 2130 protein units, thus also forming a helix of the same pitch with a diameter of 80 Å. This concept of the architecture of the virus particle is well supported by recent improvements in electron microscopic techniques (Nixon and Woods, 1960). Figure 2 shows such an electron micrograph of TMV.

The virus rod is disaggregated by many agents which disrupt secondary bonds occurring in proteins, such as heat, urea, detergents, alkali, acid, phenol and acetic acid. Some of these agents supply methods for the isolation of one or the other of the two components in native or biologically functional state while the other component is more or less degraded or denatured. Of these methods, acetic acid is the most convenient and simple for the isolation of the protein (Fraenkel-Conrat, 1957), and phenol for that of the RNA (Gierer and Schramm, 1956; Haschemeyer *et al.*, 1959).

Although the minimum molecular weight of the protein component is about 18,000, as shown later, the protein as usually obtained exhibits a molecular weight of about 100,000 (48), corresponding to an aggregation product of six molecules (Schramm *et al.*, 1955). The protein as disaggregated carefully by alkali at low ionic strength, shows a molecular weight of 36,000 (Newmark and Myers, 1957), and at high pH, or in quite dilute solution, or in 67% acetic acid, the protein shows the molecular weight of 18,000 characteristic for the monomer (Ansevin and Lauffer, 1959; Wittmann, 1959b). The electrophoretic behavior of the

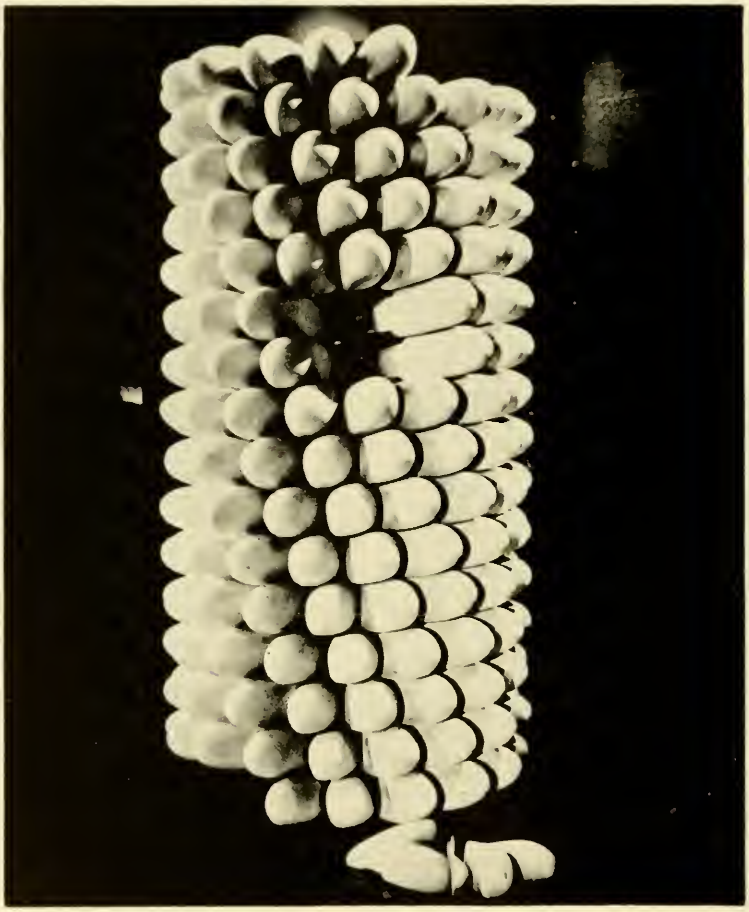


FIG. 1. Model of tobacco mosaic structure (Klug and Caspar, 1960). The model represents protein subunits (white bodies) aggregated by helical stacking and held together by RNA (black tube). Some protein units (peptide chains) have been removed to show the location of the RNA. The segment shown corresponds to about one-tenth of the total length, since one virus particle consists of about 2130 protein chains (molecular weight about 18,000), arranged in 130 helical turns.

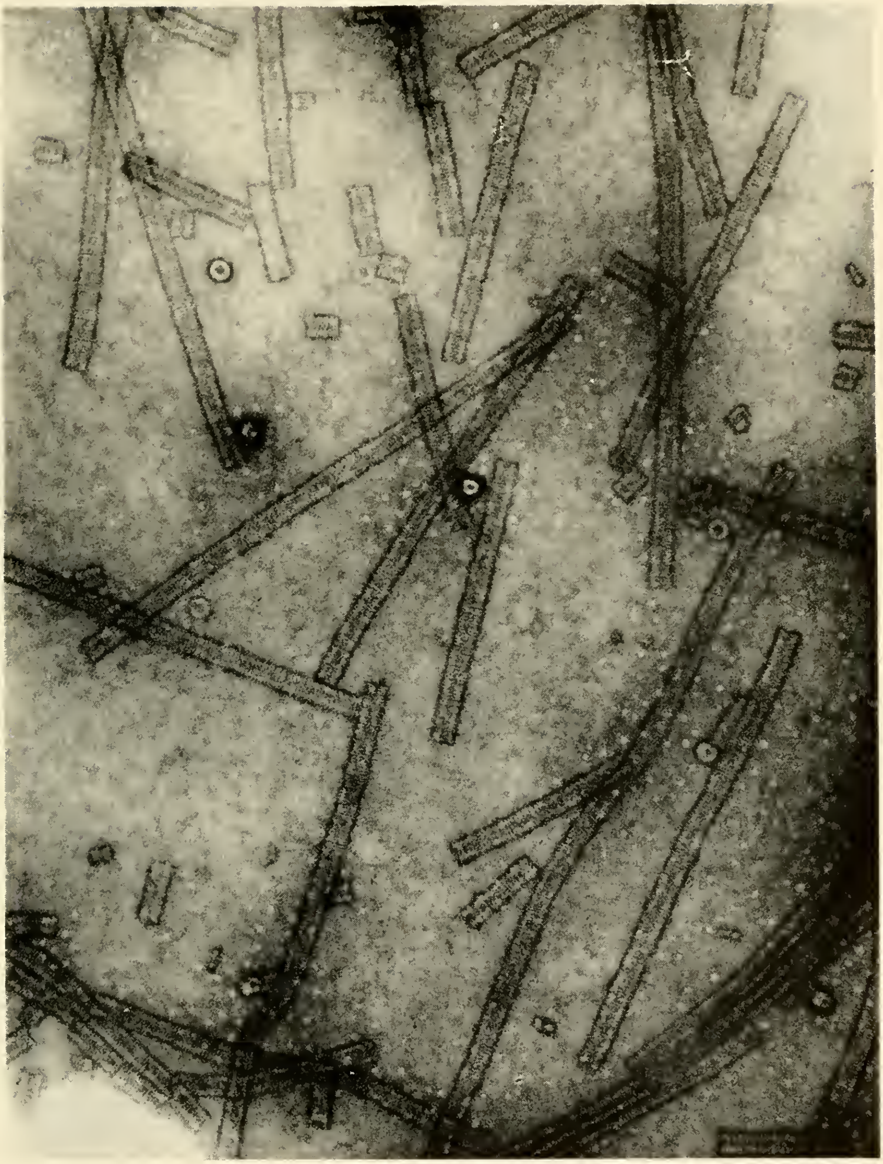


FIG. 2. Electron micrograph of TMV obtained by new technique of staining with phosphotungstate. The longest straight rod in the picture is 0.45 $\mu$ long, 50% longer than the typical length of RNA containing virus rods. The helical stacking of subunits and the central channel, here outlined by a coat of contrasting phosphotungstate solution, are clearly visible.

protein is also greatly dependent on ionic strength, pH, and temperature (Schramm and Zillig, 1955; Kramer and Wittmann, 1958).

The difficulties in obtaining physicochemical parameters for the TMV protein monomer are due to its marked tendency to aggregate, thus expressing its biological function. With increasing ionic strength and temperature, and with a pH optimum of about 5, the protein aggregates to discs, rods, and finally to rods of ever-increasing lengths, indistinguishable by all criteria, except that of irregular and, under certain conditions, excessive length, from TMV particles. These observations prove that the general architecture of the virus particle is a function only of the shape of the protein monomer, while the length of the particle, as we shall see later, is determined by the length of the polynucleotide chain.

If the shape of the protein subunit determines that of the virus, the question arises: what determines the shape of the protein? Recent studies with ribonuclease and other proteins have yielded evidence that the native conformation of a peptide chain may be a consequence of its amino acid sequence alone (White and Anfinsen, 1959). A similar conclusion appears justified in regard to the 3-dimensional structure of the TMV protein chain, on the basis of the report by Anderer (1959) that the denatured and structureless protein can readily be renatured by dialysis from 6 $M$ urea or 0.1 $N$ NaOH solution. These observations have been confirmed and extended in our laboratory. It thus appears probable that the shape of the protein monomer represents the thermodynamically most stable conformation of its peptide chain. It has not yet been demonstrated in detail what bonds account for this structure, but the occurrence of several chain segments folded into $\alpha$-helices, interspersed with non-helical areas and bends due to proline, or an accumulation of those side chains which interfere with stable helix formation appears probable (Blout *et al.*, 1960). Quite possibly the folds are stabilized by hydrophobic interactions of groups of aliphatic side chains, as well as by ionically linked and hydrogen-bonded polar residues. It must be noted that 12 of the 16 serines occur within 15 residues from both ends of the chain which are, in turn, quite low in basic and acidic residues; further, the majority of the serines, leucines, and isoleucines, as well as of the primary amide groups, occur in clusters along the chain.

Certain specific bonds have been suggested as participating in either or both, chain conformation and interchain bonding. Thus, the —SH group can be detected and titrated only in the presence of denaturating reagents, and it is particularly unusual in its behavior in the intact virus. When the virus is treated with iodine the —SH group becomes substituted. The resultant sulfenyl iodide (—SI), usually quite a labile group, is permanently stable in the virus rod but decomposes concomitantly with disaggregation of the virus. In contrast, the —SH group of the isolated protein is oxidized by iodine (Fraenkel-Conrat, 1955a). The —SH group of the virus does not react with *p*-chloromercuribenzoate, but is able to bind small mercurials (e.g., methyl mercuric nitrate) and the methyl mercury group is not dissociated from the virus by cysteine, in contrast to that bound to the isolated protein (Fraenkel-Conrat, 1959). These observations suggest that the —SH is located near one of the protein surfaces which are exposed upon dissociation of the virus, and that the masked —SH carries an H atom which can be replaced by mercurials without change in the tertiary structure. Such an —SH group may be engaged in hydrogen bonding by acting as acceptor to another X—H donor group. This would account for the unusual properties of the sulfur.

Another type of bond probably contributing to the tendency of the protein to aggregate to rod-shaped particles involves certain carboxyl groups. The disaggregation of the virus has been found to be accompanied by the release of one to two $H^+$ ions per protein unit, and, conversely, $H^+$ ions are consumed in the aggregation of the protein with or without RNA. A further interesting observation supporting the existence of this type of hydrogen atom is that about 2000 lead atoms are bound by the virus and that the introduction of lead seems to stabilize the intersubunit bonding. It has also been observed by X-ray scattering that this lead is located at two distinct cross-sectional sites in the virus particle (Caspar, 1956; Klug and Caspar, 1960). On the basis of the above experiments it has been suggested that there exists a particular H bonding involving undissociated —COOH groups which may play an important role in the build-up of the cylindrical virus structure (Fraenkel-Conrat and Narita, 1958).

The attractive forces between the nucleic acid and the protein units may be largely ionic bonds between the basic residues of the protein and the phosphate groups of the nucleic acid, but hydrogen bonds and other forces probably contribute to the stability of the coaggregate of the two components. No definite experimental approach has yet been applied to the study of these forces.

Thus, to summarize, the architecture of the TMV particle is the consequence of protein aggregation reactions at two levels: the first, the establishment of a specific chain conformation of the protein unit, and the second, the aggregation of these protein units, at first by side-to-side attachment and then by helical stacking into stable rods with a minimal protein surface area and maximal bonding of each unit to its six neighbors.

### B. AMINO ACID SEQUENCE

The amino acid composition of the TMV protein has been determined by several means. Recently the composition has been definitely established by ion exchange chromatography, using the automatic amino acid analyzer (Tsugita and Fraenkel-Conrat, 1960). Almost the same composition was deduced also by Wittmann and Braunitzer (1959) by a summation of the compositions of the separated peptides resulting from tryptic digestion of the virus protein. The two analyses agreed well except for one isoleucine residue. Nine residues of isoleucine were detected after sufficiently long hydrolysis of the protein, as well as in the sequential study at Berkeley, and that this number rather than eight was correct was later confirmed by Anderer (1962). The main problems in amino acid analysis reside now in the purity of the material, and in the rate of the liberation or decomposition of the amino acids during hydrolysis rather than in the accuracy of the analysis. Table I shows

TABLE I

Amino Acid Composition of Natural Strains of TMV[a]

| Amino acid | Group A | | | | | B | | C | D |
|---|---|---|---|---|---|---|---|---|---|
| | TMV | masked | J14D1 | YA | GA | Y-TAMV | Dahlemense | G-TAMV | HR |
| Asp | 18 | 18 | **17** | **19** | **19** | 18 | **17** | 22 | 17 |
| Thr | 16 | 16 | 16 | **17** | **17** | 17 | 17 | 19 | 14 |
| Ser | 16 | 16 | **17** | **14** | **15** | 15 | **16** | 10 | 13 |
| Glu | 16 | 16 | **15** | 16 | 16 | 19 | 19 | 16 | 22 |
| Pro | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 10 | 8 |
| Gly | 6 | 6 | 6 | 6 | **5** | 6 | 6 | 4 | 4 |
| Ala | 14 | 14 | 14 | 14 | 14 | 11 | 11 | 18 | 18 |
| Cys | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| Val | 14 | 14 | 14 | 14 | 14 | 15 | 15 | 12 | 10 |
| Ileu | 9 | 9 | 9 | **8** | **8** | 7 | 7 | 8 | 8 |
| Leu | 12 | 12 | 12 | 12 | 12 | 13 | 13 | 11 | 11 |
| Tyr | 4 | 4 | 4 | 4 | 4 | 5 | 5 | 6 | 7 |
| Phe | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 6 |
| Try | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 2 | 2 |
| Lys | 2 | 2 | **3** | 2 | 2 | 2 | 2 | 1 | 2 |
| Arg | 11 | 11 | 11 | **12** | **12** | 9 | 9 | 8 | 11 |
| Met | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 2 | 3 |
| His | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| Total | 158 | 158 | 158 | 158 | 158 | 158 | 158 | 158 | 158 |
| ± from TMV | | 0 | 2 | 3 | 3 | 8 | 8 | 17 | 17 |
| | | | | | | ± from → Y-TAMV | 1 | 18 | 17 |
| | | | | | | | | ± from G-TAMV | 16 |

[a] Numbers in boldface indicate differences from the prototype.

the amino acid composition of the protein. It should be noted that the protein lacks histidine, cystine, and methionine and has but one cysteine.

The sequential analysis was initiated at the Virus Laboratory at Berkeley in 1952. As usual in the protein field, the structural study began with end group analyses, and in particular with the finding that carboxypeptidase liberated about 2000 threonine residues per virus particle (Harris and Knight, 1952, 1955). This finding served as the first clear indication that the minimal molecular weight of the virus protein represented the subunit or peptide chain weight, a fact which was later supported by physical methods. The structure in the C-terminal region was shown by partial hydrazinolysis to be -Pro-Ala-Thr (Niu and Fraenkel-Conrat, 1955a) and the terminal hexapeptide, -Thr-Ser-Gly-Pro-Ala-Thr, was detected in chymotryptic digests soon thereafter (Niu and Fraenkel-Conrat, 1955b).

The general methods for amino terminal analysis either failed to yield any knowledge or yielded incorrect information (Braunitzer, 1956). The difficulty was explained when Narita (1958) found that an acetyl group blocked the amino-terminal end of the protein. This finding leads to interesting questions concerning the mechanism and the timing of the acetylation of the chain in relation to its biosynthesis and concerning the mode by which this feature is controlled by genetic information. This is a problem of general significance since terminal acetyl groups have since been found to occur not only in all plant viruses investigated, but also in several other proteins, e.g., ovalbumin, cytochrome c, and certain hormones.

Starting from identification of the end groups, the results of a systematic study of the amino acid sequence of the protein were reported from the Virus Laboratory in Berkeley in a series of papers (Gish *et al.*, 1958; Ramachandran and Gish, 1959; Gish, 1959, 1960, 1961; Tsugita, 1960, 1962a,b; Tsugita *et al.*, 1960). An independent study by Anderer and co-workers at the Max-Planck-Institut in Tübingen began to yield valid data in 1960 (Anderer *et al.*, 1960a). The first step in the complete analysis is the splitting of the protein into definite peptides and the separation of these peptides from one another. For this purpose trypsin was the main tool employed by both groups, since other enzymes lack the necessary specificity. Countercurrent distribution methods were selected for separation of the digested peptides at our laboratory, and more recently basic ion exchange column chromatography (Dowex 1) has been employed. Further purification was necessary to obtain most peptides in pure form. The structural studies on some of the individual peptides with as many as 41 residues are rather complicated but almost complete agreement has been reached on the structures of all but this longest one by the two research groups, and the few differences remaining between our sequences (Tsugita *et al.*, 1960, 1962a) and the recently corrected ones of Anderer (1962) will probably be soon resolved.

In order to obtain the sequential arrangement of the peptides, Wittmann (1960a) investigated the trypsin split products of several natural mutants characterized by fewer arginine and/or lysine residues than the common strain. From the comparison between the number and composition of the peptides obtained from the wild type and those from such mutants the order of some of the peptides has tentatively been established. The results were confirmed and extended by studies of the basic peptides obtained by chymotrypsin and subtilisin which supplied the overlapping sequences needed to align all tryptic peptides in a unique order (Anderer *et al.*, 1960b; Tsugita *et al.*, 1960). By these efforts the amino acid sequence of the TMV-protein has been determined almost completely in each of the two laboratories, and since the minor uncertainties do not overlap, the entire sequence can be regarded as known

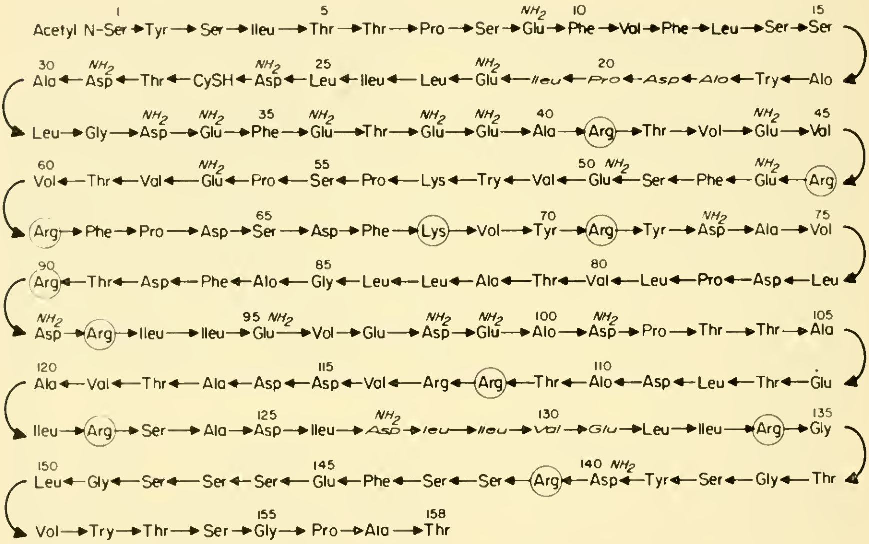

Acetyl N-Ser→Tyr→Ser→Ileu→Thr→Thr→Pro→Ser→Glu(NH2)→Phe→Val→Phe→Leu→Ser→Ser
1, 5, 10, 15
Ala←Asp(NH2)←Thr←CySH←Asp(NH2)←Leu←Ileu←Leu←Glu(NH2)←Ileu←Pro←Asp←Ala←Try←Ala
30, 25, 20
Leu→Gly→Asp(NH2)→Glu(NH2)→Phe→Glu(NH2)→Thr→Glu(NH2)→Glu(NH2)→Ala→Arg→Thr→Val→Glu(NH2)→Val
35, 40, 45
Val←Thr←Val←Glu(NH2)←Pro←Ser←Pro←Lys←Try←Val←Glu(NH2)←Ser←Phe←Glu(NH2)←Arg
60, 55, 50
Arg→Phe→Pro→Asp→Ser→Asp→Phe→Lys→Val→Tyr→Arg→Tyr→Asp(NH2)→Ala→Val
65, 70, 75
Arg←Thr←Asp←Phe←Ala←Gly←Leu←Leu←Ala←Thr←Val←Leu←Pro←Asp←Leu
90, 85, 80
Asp(NH2)→Arg→Ileu→Ileu→Glu(NH2)→Val→Glu→Asp(NH2)→Glu(NH2)→Ala→Asp(NH2)→Pro→Thr→Thr→Ala
95, 100, 105
Ala←Val←Thr←Ala←Asp←Asp←Val←Arg←Arg←Thr←Ala←Asp←Leu←Thr←Glu
120, 115, 110
Ileu→Arg→Ser→Ala→Asp→Ileu→Asp(NH2)→Leu→Ileu→Val→Glu→Leu→Ileu→Arg→Gly
125, 130, 135
Leu←Gly←Ser←Ser←Ser←Glu←Phe←Ser←Ser←Arg←Asp(NH2)←Tyr←Ser←Gly←Thr
150, 145, 140
Val→Try→Thr→Ser→Gly→Pro→Ala→Thr
155, 158

FIG. 3. Amino acid sequence of protein from common TMV. This sequence is based on all information available in February 1962. It now appears (June 1962) that one inversion (residues 25, 26) is necessary to correct the structure previously published from our laboratory (Tsugita *et al.*, 1960; Tsugita *et al.*, 1962a). About five changes have been made in the structure reported by Anderer *et al.* (1960b; Anderer, 1962). All but the positions of a few amide groups appear to be agreed upon.

(see Fig. 3). The amino acid sequence of no other polypeptide of similar size has been established at two laboratories, and thus with a similar degree of certainty. Some of the particular features of this structure have previously been mentioned in connection with the functional aspects of the protein.

## III. TMV-RNA

The amount of RNA in TMV and most other simple RNA viruses corresponds to a molecular weight of about $2 \times 10^6$. In the case of TMV the evidence from X-ray diffraction and electron microscopic analyses suggests strongly that each virus particle actually contains but one molecule of RNA of that molecular weight (Franklin *et al.*, 1957; Hart 1958). The RNA occurs as a single unbranched chain of about 6400 nucleotides running through the rod at a diameter of 80 Å and of the same helical pitch as the protein units to which it is probably bound by 3 phosphates per subunit. The length of the RNA thus limits the length of the stable virus rod to 6400/3 protein units, or a particle weight of $(2130 \times 17{,}800) + (2 \times 10^6) = 40 \times 10^6$. It should be noted that the helical path through which passes the RNA within the rod does not allow for any contact or interaction of bases (Ginoza, 1958) as indicated by the hyperchromed state of the RNA *in situ* (Bonhoeffer and Schachman, 1960).

A number of different conditions which disaggregate the protein units have been employed for the separation and isolation of the RNA. Since the RNA is more sensitive to $H^+$ and $OH^-$ ions than the protein, only those methods which proceed near neutrality yield biologically fully active RNA. Detergents (Fraenkel-Conrat and Williams, 1955; Fraenkel-Conrat *et al.*, 1957) and phenol treatment (Gierer and Schramm, 1956) fulfill this requirement, and the phenol method has found particularly wide acceptance because of its simplicity and reliability.

TMV-RNA, as isolated in such manner, appears to contain less than 1% of protein, and no other known contaminants. If the acidic clay bentonite is added in the isolation of the RNA, the resultant preparation contains less than 0.05% of protein, and from its markedly increased stability properties it is concluded that nucleases probably represent the main proteinaceous contaminant of non-bentonite-treated RNA preparations (Fraenkel-Conrat *et al.*, 1961; Singer and Fraenkel-Conrat, 1961).

Sedimentation analyses of the RNA show a main component of about 30S, which corresponds to the expected molecular weight of $2 \times 10^6$ for the complete RNA moiety of one virus particle (Gierer, 1958a,b; Boedtker, 1959). However, from 20–50% of the material shows a broad distribution of lower sedimentation rates, as it might be expected from random fragments of the predominant molecular species, which alone appears to be infectious. That the infectious molecules were single-stranded and of the length corresponding to the highest S-values was most clearly demonstrated by Gierer's study of the kinetics of enzyme inactivation of the viral RNA (1957, 1958a,b). Light-scattering and viscosity data support these conclusions. As stated, TMV-RNA, like φX174 DNA and unlike most other DNA's, exists in single-stranded form in the virus. It is also single-stranded if isolated at low ionic strength. In the presence of salts, and even at low concentrations of divalent cations, RNA undergoes marked conformational changes (Haschemeyer *et al.*, 1959; Doty e*t al.,* 1959). The drop of UV absorbance (hypochromism), and the dramatic rise in optical rotation renders RNA similar to DNA, and these effects, as well as alterations in chemical reactivity, have therefore been attributed, by analogy, to helical double-strandedness involving about 60% of the bases (Haselkorn and Doty, 1961). However, these effects are freely reversible by dilution, heat, Versene, and other agents, and it thus seems certain that the primary biological role of the RNA as a carrier of genetic information is not dependent upon such conformational features. The same is probably true also for DNA.

The chemical study of TMV-RNA began with the analysis of its component bases. The four typical bases, adenine, guanine, cytosine, and uracil, were found in the proportions of 1.13:1.00:0.77:1.10 (Knight, 1952), and no minor components such as methylpurines or pseudo-uracil have been detected (Littlefield and Dunn, 1958). Attempts to further characterize the RNA in chemical terms have followed two paths. The determination of end groups and terminal sequences may be regarded as the beginning of a systematic sequence analysis. Methods for the identification of terminal residues are available, although they have heretofore not been applied successfully to any RNA exceeding a molecular weight of 100,000 (S-RNA). Such application is dependent on the availability of RNA of the highest possible purity in terms of freedom from chain fragments or shorter oligonucleotides. For those end group methods which rely on a certain type of enzymatic digestion the absence of other nucleases from both the RNA and the enzyme employed is also of critical importance. These difficulties have largely been overcome in recent studies of this laboratory. Through the use of virus carrying much radioactive label ($4–40 \times 10^6$ cpm/mg of $C^{14}$ or $P^{32}$) it was possible to show that neither end of the chain was phosphorylated. Adenine was found to occur at both ends of the chain, and the 5′-linked terminal adenine was preceded by a pyrimidine, probably uracil (Sugiyama and Fraenkel-Conrat, 1961; Sugiyama *et al.,* 1962; Whitfeld, 1962; Fraenkel-Conrat and Singer, 1962).

The other approach is a quantitative study of the digestion products obtained from RNA with specific enzymes. Structural work along those lines may be facilitated by recent improvements in oligonucleotide fractionation methods, such as the 2-dimensional mapping procedure of Rushizky and Knight (1960), and the column chromatographic method of Staehelin *et al.* (1959).

## IV. Viral Infectivity

### A. CHEMICAL BASIS OF INFECTIVITY

It was reported in 1955 that TMV-RNA by itself had little biological activity, but that the coaggregation of native TMV protein with TMV-RNA prepared at neutrality yielded particles which were several hundred times more infective than either component (Fraenkel-Conrat and Williams, 1955). At that time it seemed that infectivity was a property only of the complete virus, be it isolated from a plant, or reconstituted *in vitro* from its components. Further studies, however, revealed that the low infectivity found in RNA preparations represented an intrinsic property of the RNA itself, and that this infectivity was affected by the protein coating only in a quantitative sense (Fraenkel-Conrat, 1956; Gierer and Schramm, 1956). This conclusion has been confirmed by others and has been supported by independent observations with other viruses including animal viruses, as reviewed by Colter (1958) and Fraenkel-Conrat (1962).

The main pieces of evidence for the intrinsic infectivity of the RNA were *(1)* that traces of ribonuclease destroyed the infectivity of the RNA, while proteases did not, *(2)* that the infectivity was sedimentable together with the RNA, and not at the high rate characteristic for the virus and *(3)* that a variety of chemical and serological tests failed to detect in the RNA significant amounts of contamination by virus protein or peptide. In addition, electron microscopy and other physical methods have been employed and all observations support the conclusion that the RNA carries the infectivity.

We may now raise the question of why the RNA has only about 0.1% of the infectivity expected from the activity of TMV. Similarly low infectivity is observed in RNA preparations from different sources and prepared by various methods. This low infectivity may be presumed to be due to the RNA being damaged in the course of purification. However, the fact that the low activity of the RNA is greatly increased upon reconstitution and then approaches that of the original virus

(Fraenkel-Conrat and Singer, 1957, 1959) disproves this explanation. On the other hand, the presence of ribonuclease in the leaves of the test plant suggests the possibility of digestion and inactivation of the RNA during the process of infection. The recent finding that the infectivity of the RNA is greatly potentiated by assay in presence of bentonite, an agent which absorbs and inactivates many nucleases, seems to support this interpretation (Singer and Fraenkel-Conrat, 1961).

### B. MECHANISM OF INFECTION

Our knowledge concerning the intracellular processes concerned with plant virus replication is as yet quite fragmentary. It seems that these viruses lack the sophisticated infection mechanisms of bacterial and animal viruses, and cellular entry of the virus appears to require a mechanical wounding, which is frequently achieved through an insect vector. Soon after entry into a viable cell the virus appears to shed its protein coat. This, in the case of TMV, requires several hours and accounts for the fact that the infection process is faster when free RNA rather than when the complete virus is used as the inoculum (Siegel *et al.*, 1957; Engler and Schramm, 1959; Fraenkel-Conrat *et al.*, 1959). Subsequently, an eclipse period of a few hours can be observed, during which the RNA is believed to enter the nucleus and initiate its replication (Wildman, 1959; Zech, 1961). The progeny RNA then seems to reenter the cytoplasm and may there initiate virus protein synthesis. While thus infective RNA appears to be formed prior to any complete virus, excess viral protein is often observed at later stages in infected cells. Thus many virus isolates contain virus-like particles lacking nucleic acid and thus void of infectivity. In the case of TMV, this so-called X-protein, first isolated by Takahashi and Ishii (1952), was more recently shown to reconstitute fully active virus when combined with TMV-RNA *in vitro* (Takahashi, 1959).

### C. CELL-FREE BIOSYNTHESIS OF TMV-PROTEIN

Nirenberg and Matthaei recently reported on an amino acid-incorporating system from *Escherichia coli* which was stimulated by the addition of macromolecular RNA (1961). The system consists of the ribosomal fraction of *E. coli*, the supernatant (containing S-RNA), an ATP generating system, and amino acids one or all of which are labeled with $C^{14}$. When the extent of amino acid incorporation into hot TCA-insoluble material was found greatly stimulated by the addition of TMV-RNA to the reaction mixture, the question arose whether the RNA might act as messenger and evoke the synthesis of TMV-protein in this cell-free system. This concept seemed to be supported when a serological relationship between the product and anti-TMV serum was detected. A detailed analysis of the labeled material produced under the influence of TMV-RNA was thereupon initiated. When unlabeled TMV-protein was added as carrier after the incubation and subsequently reisolated by column chromatography, it contained and retained a significant fraction of the counts, if TMV-RNA had been present during the reaction, but much less if another RNA or no RNA had been present. Upon tryptic degradation of the protein and separation of the typical peptides, the counts were found distributed in characteristic manner; thus, with phenylalanine or tyrosine as labeled amino acids the radioactivity was predominantly associated with the peptides known to carry the respective amino acid. These data seemed to establish rather firmly that a protein resembling TMV-protein was made by the cell-free *E. coli* enzyme system under the influence of TMV-RNA. However, one test suggested that there was something wrong with much of the product. While the carrier protein reisolated from control reaction mixtures combined with TMV-RNA in the usual way under the conditions of reconstitution, the protein carrying what was presumed to be newly synthesized TMV-protein formed very little stable virus. This strong interference with the reconstitution reaction by the small amount of "product" present may be tentatively interpreted as evidence that a good part of this protein carried a structural imperfection which prevented it from participating normally in the aggregation reaction which leads to rod formation. If one thinks of the typically aggregating protein units as analogous to bivalent antibody molecules, then the *in vitro* synthesized protein might be regarded as monovalent. Such a mental image would readily explain the marked interference of the latter with the reconstitution reaction.

To conclude this review of a number of quite recent and preliminary results, it now appears that TMV-RNA can act as messenger and evoke in a completely foreign system the synthesis of a protein which shows very great resemblance to TMV-protein in serological, chromatographic, and amino acid sequential terms, but probably differs in some as yet unknown feature necessary for the specific native conformation which enables TMV-protein to form rods (Tsugita *et al.*, 1962b).

## V. NATURAL STRAINS OF TMV

### A. SYMPTOMATOLOGY OF TMV AND ITS STRAINS

Plant virus diseases in general are either systemic, i.e., they spread throughout the plant, or they result in localized damage. In the case of TMV, different varieties of tobacco and other hosts respond typically in one or the other manner. The local lesion response observed in *Nicotiana glutinosa* and the artificial hybrid, *N. tabacum* var. *Xanthi,* nc, serves for the quantitative assay of the virus, while the spreading of the virus throughout the growing plant in *N. tabacum* is most favorable for isolation of the virus in preparative scale. Apart from the purposes of quantitation and isolation, the host response plays a crucial role in the detection and characterization of strains or mutants of the virus. Thus, many strains give characteristically small and non-spreading lesions and can in this manner be differentiated on the local lesion host. Many strains give local lesions in *N. sylvestris* and Java tobacco, varieties on which common TMV spreads systemically. Finally, the nature of the systemic response of *N. tabacum* and other varieties can be markedly different. Virus strains are known which do not show any detectable symptoms, even though they build up to high concentrations throughout the plant. Others give mild symptoms (mottling), or the typical blistered effect of the mosaic disease, which may be accompanied by more or less yellow discoloration of leaves. Finally, necrotic areas may develop upon infection with certain strains, and leaves or even the whole plant may die as a consequence of the infection. Thus, the observation of the response of several varieties of tobacco usually permits the identification of a known strain or the characterization of a new strain. The same techniques serve to test the biological homogeneity of a strain or to free it from a contaminant. Repeated passage through a local lesion host at limit dilution gives the best assurance that one is dealing with a genetically homogeneous population of virus particles. Further, the use of a differential host i.e., one which responds systemically to one strain and locally to another can rule out the simultaneous presence of both types of strains.

An additional criterion has been proposed for the differentiation of strains which also supplies data concerning the viability of such strains. This method consists in homogenizing individual lesions (8-day-old, on *N. tabacum* var. *Xanthi,* nc) and testing these homogenates after appropriate dilution on new *Xanthi* plants. The number of lesions which can be evoked by one parental lesion has been termed productivity and was found to vary over tenfold for different natural strains (Veldee and Fraenkel-Conrat, 1962).

Since mechanical wounding is required to initiate infection, the virus solution is usually rubbed onto the surface of the leaves in presence of an abrasive (i.e., carborundum). Any trace of virus accidentally deposited on the leaf surface at any time prior to this rubbing through transmittal by agents such as the nozzle of a water hose can upon rubbing with an abrasive evoke lesions which then are not due to the strain used in the inoculum, obviously a serious source of error.

For the quantitative testing for infective virus particles solutions containing on the order of 1 to $10 \times 10^{-6}$ mg virus/ml are rubbed onto half-leaves of the *Xanthi*, nc variety, the most sensitive test plant, and the resultant lesion number compared to that given by a known virus solution, preferably applied to the opposite half-leaves of the same plants. Over the range of 5 to 50 lesions per half-leaf, the response is approximately proportional to the virus concentration, but in absolute terms of multiplicity plant virus infection is very inefficient. The fact that it takes somewhere between $10^3$ and $10^5$ particles to produce a lesion or to initiate a systemic disease has been cause for some concern as to whether the population of virus particles is uniformly infective. However, most workers in the field tend to attribute this inefficiency of plant virus infection to the crude mechanics of the wounding and infecting process. The fact that most of the virus which is applied and rubbed onto the leaf can be washed off a few seconds later without decreasing the resultant lesion count seems to support this interpretation.

### B. COMPOSITION OF NATURAL STRAINS

Analytical comparisons of natural strains of TMV were initiated by Knight and Stanley in 1941 and the first complete amino acid analyses appeared in 1947 (Knight). The strains originally investigated, the masked (M), J14D1 (JD), yellow aucuba (YA), green aucuba (GA), Holmes ribgrass (HR), and cucumber viruses 3 and 4 (CV3 and CV4) were obtained from field isolates from various diseased plants, and all but the last two were classified as strains of TMV on the basis of physicochemical, serological, and biological tests.[2] Later Black and Knight (1953) investigated two series of more closely related TMV strains, in part derived under such greenhouse conditions from one another, that single-step mutations may be presumed.

Analyses of the proteins of natural strains were also reported from the Max Planck Institut für Biologie (Aach, 1958; Wittmann, 1960b), and nucleotide analysis were obtained by Black and Knight (1953), as well as by Markham and Smith (1950).

The base analyses of the RNA of different strains have revealed no differences in composition, quite in contrast to the RNA of different viruses, which was found to vary widely (Knight, 1952, 1959). It appears to be a general phenomenon that within a family of virus strains the RNA composition seems alike, but it must be remembered that each base occurs about 1500 times per molecule and that the error of the analytical method is at best ±30 residues.

---

[2] Cucumber viruses 3 and 4 were classed as TMV strains on the basis of physicochemical relationship. When chemical studies (RNA composition, end groups, cysteine content, etc.) set them sharply apart from all bona fide strains of TMV, Knight suggested reclassification (Knight, 1955).

In contrast to the uniformity in RNA composition, the amino acid make-up of most TMV strains was found to vary. Table I summarizes the results of recent analyses (Tsugita, 1962a; Knight *et al.*, 1962) which in general bear out the earlier data. These analyses show that TMV strains can conveniently be classified in the following analytical groups: A. One group of many strains which are either identical with TMV, or quite similar to it, showing 1–3 amino acid residues replaced by others, and always lacking histidine and methionine. To this group belong most of the seemingly one-step mutants detected under greenhouse conditions. B. Another group, exemplified by Y-TAMV, showing a large change in composition including one methionine and *C*-terminal serine (Knight *et al.*, 1962). Apparently the Dahlemense strain analyzed by Wittmann (1960b) belongs to this group. C. G-TAMV (again quite different, with two methionines). D. The Holmes ribgrass strain [again different, including three methionines and 1 histidine (Tsugita, 1962a)].

All strains, even those showing a very different composition, appear to contain the same number of residues, namely, 158 per peptide chain. Almost all strains studied in our laboratory which differ in composition from TMV give local lesions on *N. sylvestris*, rather than the systemic disease, but no other correlations between composition and symptomatology were detected. Thus, strains of the same group gave no symptoms on *N. tabacum*, or killed the plant; and some of the symptoms due to strains of all groups could resemble one another on one or the other host.

Recent studies of representatives of these classes of TMV strains have not confined themselves to amino acid composition. For the purpose of locating the amino acid exchanges on the peptide map, analyses of the amino acid composition of the peptides derived from trypsin digest of the protein, as well as preliminary studies of their amino acid sequences, have been initiated. Figure 4 shows comparative aspects of the tentative structure of parts of these proteins (Tsugita, 1962a). These data suggest some interesting generalizations. Although the difference in net composition between two strains may appear relatively small (e.g., 8 exchanges for Y-TAMV as compared to TMV), the difference in peptide composition is much greater, involving in the pair used as example more than 30 exchanges, i.e., 20% of the total amino acids. On the other hand, some particular peptide sequences such as Arg.Asp-$NH_2$. Arg (90–92), Val.Tyr.Arg (69–71), and Arg.Arg.Val.Asp.Asp.Ala.Thr.

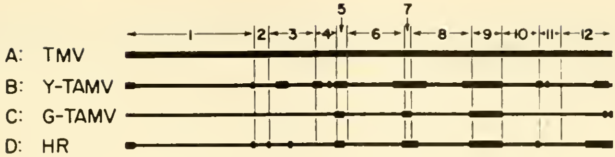


Fig. 4. Location of differences in protein structure of natural strains of TMV. The peptide chain of common TMV, subdivided at the sites of trypsin attack into 12 peptides is illustrated (A), as well as segments of this structure in which differences have been observed (*thin line*) in strains of groups B, C, and D (see text). It may be noted that peptides 5, 7, and 9 are the same throughout.

Val.Ala.Ileu.Arg (112–122) are not changed throughout the four groups, which indicates that these sequences might play an important role in the biological activity of the protein molecule. These data are as yet quite fragmentary, but it is believed that they will ultimately supply the basis for important genetic and protein structural considerations.

### C. MIXED RECONSTITUTION

Although DNA has long been regarded as the sole genetic material on the molecular level, it has been unequivocally proven in recent years that RNA fulfills this function for the RNA viruses. This was first clearly demonstrated by means of the reconstitution technique (Fraenkel-Conrat and Singer, 1957). Three kinds of natural strains and the normal TMV strain were chosen for this purpose. These strains could be distinguished from the normal strain by either their symptoms on plants or their amino acid compositions, or both. Both the protein and the nucleic acid components were isolated from common TMV and from the natural strains, and the various possible combinations of pairs of these were reconstituted and tested. For instance, with the HR strain the following couples of reconstitution experiments were made:

HR-protein + HR-nucleic acid (a)
HR-protein + TMV-nucleic acid (b)
TMV-protein + HR-nucleic acid (c)
TMV-protein + TMV-nucleic acid (d)

Each mixed reconstituted virus was allowed to infect tobacco plants and the resultant symptoms were observed, and the amino acid composition of the progeny protein determined. The results showed clearly that the nucleic acid determined both the symptoms and the amino acid composition, so that the progeny from (a) and (c) seemed to be HR and those from (b) and (d) seemed to be TMV both chemically and biologically. In other words, the nucleic acid of TMV was experimentally proved to be the sole genetic material carrying the complete information for its own structure as well as for the structure of its homologous protein, undisturbed by the presence of heterologous protein in the infecting particle (see Fig. 5).

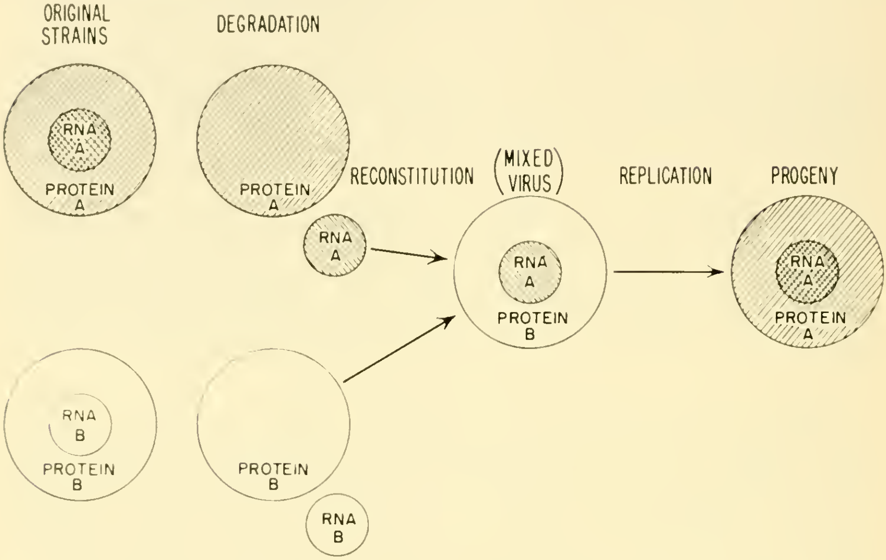

ORIGINAL STRAINS
DEGRADATION
RNA A
PROTEIN A
PROTEIN A
RNA A
RECONSTITUTION
(MIXED VIRUS)
RNA A
PROTEIN B
REPLICATION
PROGENY
RNA A
PROTEIN A
RNA B
PROTEIN B
PROTEIN B
RNA B

Fig. 5. Mixed reconstitution. The identity of the progeny protein, as compared to the two parental strains, symbolized as A and B, has been established by serological tests and amino acid analyses.

## D. STRUCTURAL DIFFERENCES IN THE RNA OF DIFFERENT STRAINS

While differences in the amino acid composition of viral strains have long been known to exist, attempts to detect differences in nucleotide composition have been unsuccessful. However, as previously pointed out, an analytical error as low as ±0.5% would give the composition of TMV-RNA with an accuracy of only ±6 to 9 residues for each base. Therefore, one cannot conclude from the available data that the strains of a given virus are identical in composition, but only that they resemble one another. Actually, present knowledge of the genetic capability of the RNA strongly suggests that there must exist structural differences between the RNA of viral strains. Such differences could conceivably be due to base interchange and thus become evident only from sequential analyses rather than from the composition of the RNA.

Experimental searches for structural difference of the RNA from strains have only recently been begun. End group analyses of four strains have shown these to have the same terminal adenine residues as the common TMV (Sugiyama and Fraenkel-Conrat, 1962), even though the proteins of two of these show differences in the N-terminal region of the peptide chain. Further information has been obtained by Rushizky and Knight (1960) through comparative studies of enzymatic digests of various strain nucleic acids. Pancreatic ribonuclease, which splits only next to pyrimidines, has been found useful for this purpose. By the 2-dimensional mapping procedure (electrophoresis and chromatography) many of the smaller digestion products can be separated and quantitated (cytidylic and uridylic acid, as well as di-, tri-, and tetranucleotides composed of purines and terminal 5′-linked pyrimidines). These studies have revealed definite differences in the frequency of recurrence of some oligonucleotide sequences in natural strains not closely related to TMV.

## VI. Chemically Produced Strains

### A. MODIFICATION OF THE RNA

Ever since the early days of their isolation the viruses have been subjected to various reagents or conditions which modified their structure (Stanley, 1936). Only more recently has it become evident that most of those agents which inactivated viruses did so by force of their affinity for, and reaction with, the viral RNA. The three main classes of modifying agents are *(1)* physical agents, such as ultraviolet light, X-rays, heat, and sonic oscillations; *(2)* the biological approach of allowing the virus to replicate in the presence of base analogs which then tend to become incorporated into the virus structure; and *(3)* a host of chemical reagents which affect the amino groups or substitute the rings of the purines or pyrimidines.

This subject has been critically reviewed by Schuster (1960), and will not be systematically dealt with here. The modification reactions usually lead to inactivation of the viruses and thus are of some interest as antiviral agents, and for the purposes of vaccine preparation. In connection with the present chapter, the main interest in these agents resides in their ability to produce mutations and they will be discussed primarily from this viewpoint.

The effects of the physical agents are varied, ranging from chain breakage to modification of the pyrimidine rings, but none of these reactions appears to lead to the frequent production of mutants. The same is true for the incorporation of base analogs. Azaguanine appears to replace some guanine residues (Matthews and Smith, 1955) and 2-thiouracil (Jeener and Rosseels, 1953; Mandel *et al.*, 1957) as well as 5-fluorouracil (Gordon and Staehelin, 1958) replace uracil, the latter up to 50%. These events are accompanied by more or less extensive inactivation (no inactivation with fluorouracil), but no mutants have been observed as a consequence of these reactions.

The chemical modification reactions are performed directly with the virus, or preferably with the isolated RNA. Formaldehyde adds to or condenses with the amino groups of the RNA (Fraenkel-Conrat, 1955b). and is a relatively inefficient inactivating agent, probably because of the reversibility of the reaction (Staehelin, 1958). Glyoxal and similar compounds have a particular affinity for guanine residues, and are better inactivating agents (Staehelin, 1959a,b). No mutants have been obtained as the result of aldehyde treatment. However, it must be stressed that the only technique which renders the detection of relatively infrequent mutants a not too laborious task, the differential host method, has not been systematically nor frequently applied to RNA modified by the agents discussed in this and the preceding paragraph.

The next class of agents to be discussed are the alkylating agents. These vary greatly in their affinity for RNA, with mustard gas being the most reactive and iodoacetate the least reactive of the series that was investigated (Fraenkel-Conrat, 1961). It appears probable that the introduction of one or two alkyl groups per mole of RNA (6400 nucleotides) causes inactivation, regardless of the reagent used. The most readily alkylatable sites are the guanine and adenine residues. When various alkylated RNA preparations were surveyed for the presence of mutants the methylated derivates were found to contain significant numbers of mutated molecules among the survivors, but all other bigger substituents appeared to be not clearly mutagenic under the same test conditions.

One other agent, *N*-bromosuccinimide, which seems to introduce bromine predominantly into the pyrimidine rings, was reported in preliminary manner to be mutagenic to a similar extent as methylation (Fraenkel-Conrat, 1961). Hydroxylamine and similar agents which attack the pyrimidine bases have also been reported to effect mutations (Schuster, 1961).

The reaction which has proven particularly fruitful in the field of virus mutagenesis has been the deamination reaction by means of nitrous acid. Schuster and Schramm (1958) and Vielmetter and Schuster (1960) established that the purine bases were deaminated at similar rates and slightly faster than cytosine in RNA, and that one hit per 6000 nucleotides had a 50% chance of inactivating the molecule. Gierer and Mundry (1958) subsequently showed that many of the non-lethal hits were mutagenic. The single-hit nature of the curve relating mutation frequency to reaction time showed that a single deamination could result in a mutation. Further work from other laboratories confirmed these observations, and it now appears that with average hit numbers greater than 1 or 2, almost all deamination survivors are mutants, though only a fraction of these give the differential host test (i.e., local lesions in *N. sylvestris* or Java tobacco). It has been established beyond doubt that deamination is per se truly mutagenic, and that selection does not play a significant role in this phenomenon (Mundry, 1959).

Having discussed the reactions which were found to effect mutations in RNA, a consideration of the presumed nature of the mutagenic event, in terms of RNA structure, appears advisable. The reaction most frequently employed is that with nitrous acid. In the case of DNA, evidence was adduced that the two reactions involving the 6-amino position (adenine → hypoxanthine, and cytosine → uracil) can be mutagenic, but that deamination of the guanine is only lethal (Vielmetter and Schuster, 1960). This is interpreted in terms of the current concepts of the replication mechanism based on the Watson-Crick scheme of base pairing. In RNA, and only in RNA, is nitrous acid an astonishingly effective mutagen, leading to mutation of almost all survivors. This has been attributed to the obligatory mutagenic effect of the deamination of cytosine since this leads to another genetically fully competent symbol in the RNA language but not in the DNA language, i.e., uracil (Fraenkel-Conrat, 1961; Tsugita, 1961). In contrast, the deamination of adenine leads to a base equally unnatural to RNA and to DNA. Thus, nitrous acid acting on RNA is the only direct mutagenic reaction known. Its effect contrasts with the various indirect mutagenic reactions, such as the deamination of the purines, or of cytosine in DNA, the methylation of the purines, the action of *N*-bromosuccinimide or hydroxylamine on the pyrimidines, the biological incorporation of base analogs into DNA, and others. All of the indirect mutagens are believed to increase the probability of a misreading of the coded information, but only the deamination of cytosine is believed to directly alter that information.

### B. PROTEIN COMPOSITION OF CHEMICALLY PRODUCED MUTANTS

When mutants became available through chemical modification of the RNA, research programs were initiated, both at our laboratory and at the Max-Planck-Institut für Biologie, to determine the composition and the structure of the proteins of these mutant strains. Since the amino acid sequence of the wild type protein was known, and analytical methods for the detection and localization of single amino acid exchanges were available, there was reason for hope that this approach would reveal significant information concerning the relationship between the structure of the template RNA and the protein product.

The methods employed by us were briefly as follows. The RNA was subjected to one of a variety of chemical agents, then freed of the reagent, and reconstituted with untreated protein. Only rarely was the complete virus subjected to the chemical treatment. As previously stated, only methylation and bromination yielded significant numbers of mutants while deamination mutated almost all survivors. The extent of chemical modification was gaged from the inactivation, 37, 14, 5, and 2% residual infectivity being regarded as indicating 1, 2, 3, and 4 inactivating hits per average molecule ($e^{-1} = 0.37$). Previous studies had demonstrated that each chemical modification, as ascertained analytically, had at least a 50% chance of being an inactivating event.

Two techniques were employed for the detection of mutants. Differential host mutants, as defined by Siegel (1960), are those which are detected as local lesions on *N. sylvestris*, a systemic host for the wild type virus. Such local lesions were individually excised and repeatedly transferred to the same host until they produced local lesions unaccompanied by systemic symptoms and thus appeared to have been freed from contaminating wild type virus particles. Symptom mutants are detected by inoculating the reconstituted RNA onto a local lesion host, *Xanthi*, nc, and then transferring homogenates of randomly selected (or all) individual lesions to *N. tabacum*, *N. sylvestris*, and again to *Xanthi*. Mutants are distinguished by the appearance of unusual numbers or shapes of lesions on *Xanthi*, nc, unusually mild or severe symptoms on the systemic hosts, or other odd symptoms, singly or combined. To verify the genetically stable nature of any such symptoms, the virus was usually passed at least once through a local lesion and again onto that host which showed a distinctive reaction. Stable mutants were then transferred to 24–72 plants of *N. tabacum* and virus harvested from these after 21 days. The symptomatology of the mutant progeny was again verified by plant tests.

The mutant virus was split with acetic acid, and its protein was isolated. Samples of 4–5 mg of protein were hydrolyzed with 6 $N$ HCl at 108 ± 1°C for 24 and 72 hours. It is important that the acid be evaporated in a few minutes to avoid losses in serine and glutamic acid. One milliliter of the residue, redissolved to about 1 mg/ml, was applied to the amino acid analyzer (Tsugita and Fraenkel-Conrat, 1962).

The question arises whether direct analysis of hydrolyzates of a protein of molecular weight 18,000 can yield amino acid composition data sufficiently accurate to detect a single amino acid change. Since the most frequently occurring amino acid in TMV is aspartic acid with 18 residues per mole, an analytical error of ±2.5% would represent the limit for the accurate determination of this integral number. An accuracy better than this (less than ±2%) can generally be obtained with the analyzer and repeat analyses performed on each of two hydrolyzates can supply enough data to establish a reasonable degree of certainty.

In contrast to our primary use of direct analysis (Tsugita and Fraenkel-Conrat, 1960) Wittmann continues to determine the composition of strains by the summation of the analyses of the peptide fragments obtained by trypsin digestion. That approach is less dependent on high accuracy in the analytical techniques, but the requirement that each peptide be obtained in a state of high purity makes this method laborious and expensive. The fact that the analytical results for common TMV obtained with the whole protein are in agreement with those obtained by peptide summation (see Section II,B) validates both methods; the direct method, being the simpler, thus appears preferable for the primary survey of a large number of mutant proteins.

For mutants showing differences in composition from the wild type the question arises as to which particular amino acids are exchanged along the peptide sequence of the protein. Obviously it is important to establish the nature and the location of an exchange of amino acids, including an identification of glutamine and asparagine as contrasted to glutamic and aspartic acid.

For this purpose 20–300 mg of the proteins were digested by trypsin in the usual manner. From the digest the isoelectrically insoluble peptide was separated, and the soluble peptides were fractionated on Dowex $1 \times 2$, using a volatile solvent system similar to that used by Wittmann and Braunitzer (1959), as shown in the caption of Fig. 6. A typical peptide elution pattern is shown in that figure. A few of the peaks appeared to consist of a single peptide as shown by rechromatography and analytical composition. Others were mixtures, but upon analysis of aliquots from the rising or falling edge of the curve the composition of the pure components was often closely approximated. For preparative isolation and definitive analyses such mixtures were rechromatographed. For the purpose of amino acid analysis of peptides the 50-cm and 5-cm columns of the analyzer were usually used rather than the 150-cm and 15-cm columns. This procedure has the advantage that the time for analysis is but one-third that of the normal procedure and that only one-tenth as large a sample is required (0.01–0.3 $\mu$moles), without much loss in accuracy (±5%) (Tsugita, 1962a).

Any peptide showing a change in composition from that of the corresponding peptide of the normal strain was subjected to the usual variety of techniques of protein chemistry, i.e., digestion by proteinases or peptidases and chemical methods including dinitrophenylation, phenylisothiocyanation, and hydrazinolysis. Thus, the localization of the specific exchange was usually possible. Certain changes, involving residues near the *N*-acetyl end or the C-terminus, could be ascertained by application of similar methods to the intact protein.

We shall now proceed to consider the results of these studies (see Table II). The first important observation is that many mutants (about 50% of those analyzed) did not differ from common TMV in their amino acid composition. Thus, the original report of Wittmann (1959a) concerning a mutant which showed the same composition as common TMV has been confirmed.

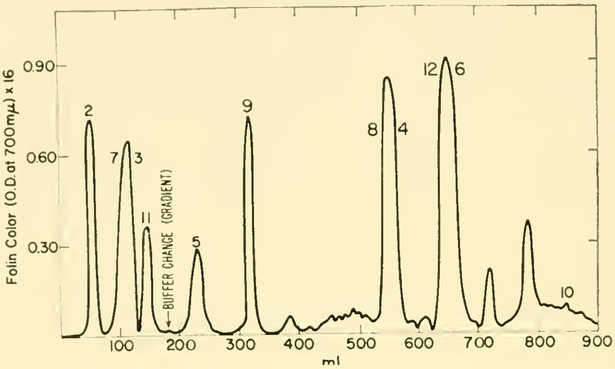


FIG. 6. Chromatographic separation of peptides in soluble fraction of tryptic digest of TMV protein. *Buffer system:*

1. 40 ml collidine + 40 ml pyridine + 1.5 ml acetic acid (pH 8.2) (180 ml).

2. Gradient (by Autograde) (Technicon Chromatography Co., New York) Chambers *1* and *2*: 40 ml collidine + 40 ml pyridine + 8.0 ml acetic acid (pH 7.3); *3*: 0.1 *N* acetic acid; *4*: 0.5 *N* acetic acid; *5*: 0.5 *N* acetic acid; *6*: 1 *N* acetic acid; *7*: 2 *N* acetic acid; *8* and *9*: glacial acetic acid (each chamber contains 120 ml).

Column, 0.9 × 150 cm, Dowex 1 × 2 (200–400 mesh); flow rate 30 ml/hour by Mini pump (Milton Roy Co.); fraction, 3.3 ml; 80 mg protein.

The numbers identify the main component of each peak with the sequential peptide number on the polypeptide chain; 1 is not present because it is the N-terminal I-peptide removed by centrifugation prior to chromatography.

Another class of mutants was observed which showed the replacement of 1–3 amino acids. In no case was the total number of residues found to be changed. Certain amino acids were observed to be frequently substituted, and frequently by a particular one; others have not as yet been found to be replaced. Although our studies covered mutants obtained with three different types of reactions (deamination, alkylation, and bromination), the extent and the nature of the observed changes were similar. Thus, 1 deamination, 4 bromination, and 3 methylation mutants showed the replacement of 1 proline by a leucine, which in-

TABLE II

Amino Acid Exchanges in Chemical Mutants of Common TMV[a]

| Mutagen | Mutant No. | Amino acid exchanges | Location |
|---|---|---|---|
| $HNO_2$ | 273 | Asp[b] → Ser | Peptide #1 |
| $HNO_2$ | 282 | Asp[b] → Ala | Peptide #10 |
| | | Thr → Ser | Peptide #1 |
| $HNO_2$ | 171 | Pro → Leu | Residue #156 |
| | | Asp[b] → Ala | Peptide #10 |
| | | Thr → Ser | Peptide #1 |
| $HNO_2$ | 332 | Pro → Leu | Peptide #1 |
| $HNO_2$ | 220 | Ser → Phe | Residue #138 |
| $HNO_2$ | 329 | Ser → Phe | Residue #138 |
| | | Ser → Phe | Peptide #12 |
| $HNO_2$ | 237 | Arg → Gly | Residue #61 |
| | | Arg → Gly | Residue #134 |
| $HNO_2$ | 321B[c] | Arg → Gly | Residue #61 |
| | | Arg → Gly | Residue #134 |
| | | Arg → Gly | Residue #122 |
| $HNO_2$ | 284 | Glu → Gly | Residue #97 |
| | | Arg → Lys | |
| $HNO_2$ | 328 | 16 amino acids exchanged in many peptides; composition identical with G-TAMV (see Table I) | |
| $HNO_2$ | 262 | | |
| NBSI | 218 | Pro → Leu | Peptide #1 |
| NBSI | 233[d] | Pro → Leu | Peptide #1 |
| NBSI | 207 | Pro → Leu | Peptide #1 |
| NBSI | 235 | Pro → Leu | Peptide #1 |
| NBSI | 187[d] | Arg → Gly | Residue #46 |
| | | Asp[b] → Ser | Peptide #1 |
| NBSI | 326 | Asp[b] → Ser | Peptide #1 |
| | | Asp[b] → Ser | Peptide #1 |
| | | Asp[b] → Ser | |

*Table continued*

TABLE II (*Continued*)

| Mutagen | Mutant No. | Amino acid exchanges | Location |
|---|---|---|---|
| NBSI | 330 | Ileu → Thr | |
| NBSI<br>NBSI | 223<br>206 | 17 exchanges, similar to G-TAMV, but for 1 Ala → Gly | |
| NBSI | 331 | 7 exchanges, similar to Y-TAMV (Table I), but for 1 Leu → Ileu | |
| DMS | 214 | Pro → Leu | Peptide #1 |
| DMS | 278 | Pro → Leu | Peptide #1 |
| DMS | 176 | Pro → Leu | Peptide #1 |
| DMS | 215 | Ser → Phe | Residue #138 |
| DMS | 178 | Arg → Gly | |
| P.O.[e] | 249 | 16 amino acid exchanges in many peptides; composition identical with G-TAMV | |

[a] Besides these 27 strains showing differences from common TMV, which in most instances were confirmed also by the analyses of progeny isolates, 29 other strains showed no detectable differences in composition from the parental strain.

[b] Not known whether aspartic acid or asparagine residue.

[c] Secondary change from strain 237.

[d] Secondary change.

[e] Propylene oxide.

volved not the same near *C*-terminal proline previously reported to be replaced in 1 nitrous acid mutant (Tsugita and Fraenkel-Conrat, 1960); and 1 particular serine (#138) was found replaced by phenylalanine in 1 methylation and 2 deamination mutants. These data are summarized on Table II.

A survey of the location of changes on the peptide map indicates that these occur at sites scattered over the entire sequence. Where multiple changes were observed, these were not in vicinal positions. While the total number of located changes is too small to justify any far-reaching conclusions, it would appear from the data illustrated in Fig. 7 that the distal parts of the peptide sequence are more prone to change than the central ones. It must be stressed, however, that this may have no genetic cause but be due to selection, in that mutation resulting in the exchange of specific amino acids might not give a functional or stable protein, and thus produce either no virus progeny, or too little material for chemical analysis.

We have discussed one class of mutants showing the same amino acid composition as common TMV, and another group showing a limited number of alterations. Finally, another class was observed at our laboratory which was very markedly different from the parental strain, differing by 8 or by 16–17 net amino acid exchanges, and probably in many more sites on the molecule. Representatives of this group closely resembled two natural strains previously described (groups B and C in Table I), and concern arose that such virus isolates were actually derived from contaminating lesions. The finding that three of the five members of this group clearly differed from the natural strains by the replacement of 1 or 2 amino acids, in one instance involving an increase in glycine content by 25% seemed to rule out this interpretation. Concerning the origin of these strains, they were derived from random lesions on *N. sylvestris* resulting from the same reaction mixtures (deamination, bromination, and alkylation), which also yielded strains showing from 0 to 2 amino acid replacements. The average number of hits per particle, as judged from the level of inactivation was 3–6. Thus, no reason becomes evident for the appearance of a new family of mutants

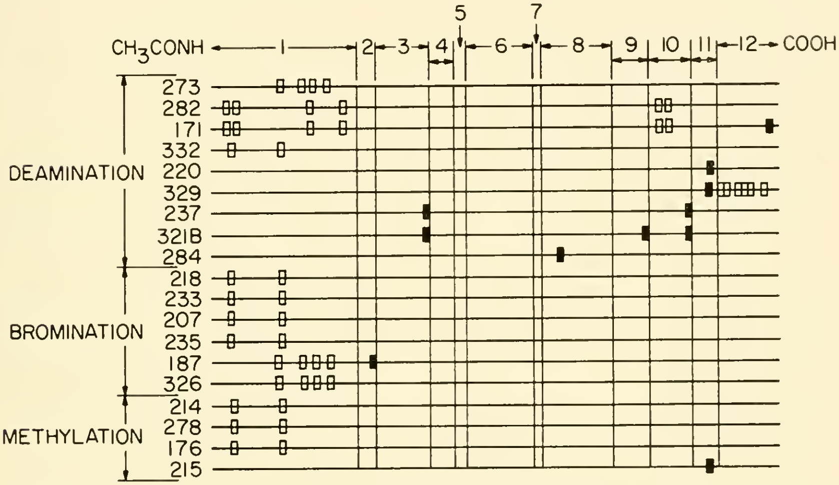


FIG. 7. Location of amino acid exchanges in chemically evoked mutants. Each horizontal line represents the peptide chain of a mutant, subdivided into tryptic peptides as in Fig. 4. Solid rectangles indicate definite location of a change, open rectangles possible alternate sites of a single change in a specific sequential peptide (#1–12). Only in strain 326 are 2 of the 4 potential aspartic acids of peptide #1 transformed to serines.

showing a composition so grossly unlike the parental strain and most of the mutants resulting from it.

While no explanation can be offered for these observations, a review of what we know concerning the composition of various natural strains may supply some basis of understanding. As previously discussed, natural TMV strains are known to occur in groups: the M strain is indistinguishable from the common strain, and YA, GA, and J14D1 differ by 2 to 3 residues. Then suddenly we are faced with several extremely different strains, bearing little relationship to TMV and in only a few instances resembling one another. Yet no strains of intermediate compositions are known, although this would be expected if the latter arose by a series of 20 or more mutational events from common TMV. Thus, the situation is similar for both series, that of the naturally observed and of the chemically evoked strains, even though this fact does not facilitate an understanding of the mechanism of either. The possibility must naturally be considered that any isolated mutant is actually due to a spontaneous mutation, rather than to the chemical treatment employed. This possibility cannot be excluded in any individual case, and therefore no conclusions should be based on single isolates. Yet, obviously, it would be hardly less surprising to find spontaneous mutation to lead to such marked changes. Further the infrequency of detection of spontaneous mutation makes this appear unlikely on statistical grounds, since the unexpected observation of gross changes in protein composition upon treatment with three different mutagens were made repeatedly.

Finally, there exists the possibility that the various chemical agents through their inactivating action might select for a more stable strain or family of strains present in traces in common TMV preparations. This might seemingly account for the six observed strains of unusual composition, all of which resulted from preparations inactivated to 95% or more. Nevertheless, such an explanation appears improbably because common TMV generally appears more stable and more viable than its strains. Furthermore, the fact that the RNA rather than the whole virus was subjected to the inactivating reactions makes this interpretation even more unlikely, for it appears very improbable that the RNA differs greatly in stability from one strain to another.

In the course of our chemical analyses of many mutants, these were repeatedly passed through single lesion hosts and through large groups of *N. tabacum*. Occasionally, and particularly in hot weather, the virus progeny so isolated was not identical in composition or symptomatology with the parental strain. At times, altered symptoms were detected on part of the inoculated plants and these plants were found to contain the chemically different strains. Of the six instances in which such phenomena were observed, four showed a single "secondary" amino acid exchange. This could be due to some heterogeneity in the primary mutant, with a minor component (which then could be attributed to the original modification reaction) being favored by the test conditions. On the other hand, these observations may also signify that some of the chemically evoked mutants are rather unstable and readily undergo additional changes due to spontaneous mutations in the course of replication. The corollary of this hypothesis would be that the spontaneous mutation frequency is much higher for RNA than generally presumed, but that this fact is usually concealed by the lesser viability of the vast majority of the mutants as compared to the parental strain. If this were the case, then the observed mutants showing very different composition from the parental virus might be attributed to a series of spontaneous mutational events within a low-viability population, leading ultimately to a stable or rather a viable virus which then, not too surprisingly, would tend to resemble a known natural strain. In that case no direct relationship between the mutagenic reaction and the final stable strain would be expected.

These considerations probably do not apply to the strains showing only one or two exchanges, and yet among these also the same amino acid replacements were frequently observed to recur, regardless of the mutagenic reagent used.

In conclusion, then, no evident correlation was detected between the presence and extent of amino acid exchanges and either the nature of the modifying reaction or its intensity. The only clear correlation which was noted in this work was that amino acid changes were detected in almost all mutants giving local lesions on *N. sylvestris*, and in only one mutant not showing this distinct biological property. It is this recognition which enabled us to select more promising mutants for chemical study, and thus to obtain relatively more data on amino acid exchanges than could be obtained by random analysis of mutants irrespective of the nature of their symptom alterations.

Parallel to our studies Wittmann analyzed a considerable number of nitrous acid mutants (1960a, 1961). His techniques differ from our in the following regards. He modified the virus directly, rather than the RNA, and has used only nitrous acid as the mutagen. As starting material he frequently employed a strain (A-14), rather than common TMV. For the detection of mutants, Wittmann relied on the transfer of single primary lesions from *Xanthi* to *N. tabaccum* (symtom mutants) avoiding selection on *N. sylvestris*. The reason for this was his desire to minimize the danger of secondary mutations in the course of repeated transfer. After isolation of the progeny protein, Wittmann directly proceeded to tryptic splitting, and analyzed the resultant peptides separately. He has never reported in detail how he purified each peptide and it is our experience that some, and particularly the I-peptide (41 residues), can be freed only with difficulty from traces of adsorbed peptides. The possibility thus exists that a seeming difference in the composition of the I-peptide is due to this technical difficulty. The results reported by Wittmann are in part in line with our observations, but show some unexpected differences. His observed proportion of mutants of unchanged composition (70%) is higher than ours, as would be expected from the fact that he relied on random sampling only, while we frequently used a method which selects for protein changes.

A brief discussion of this group of mutants of unchanged protein composition might here be interjected. These mutants are of considerable interest since their presence clearly signifies that part of the RNA carries information which is not related to the structure of the virus coat protein. This is not a surprising fact, but had nevertheless not been previously demonstrated, and it is important to evaluate the validity of the evidence on which it is based. It would seem that the failure to detect a difference between two proteins might well be due to methodological causes. Thus, interchanges of two amino acids within one peptide, and translocation of amide groups have not been ruled out by the methods used, although the latter occurrence would probably be revealed upon chromatographic separation of the peptides. In view of these possibilities, the failure to detect an actual difference is a definite possibility for any one strain protein, even though it is not very probable. However, the probability of this having occurred 75 times in one laboratory, and about 30 times in another is as good as nil, and thus the conclusion that different strains may share the same protein appears on statistical grounds firmly established.

Besides this group, Wittmann also observed mutants showing 1 or 2 amino acid replacements, but none showing the many alterations observed six times in our laboratory. His exchanges were also often the same, but oddly enough his most frequent exchanges did not coincide with those of our laboratory. Yet in certain instances (e.g., serine #138 → phenylalanine) the same exchanges were detected in both laboratories, an occurrence which would be statistically most improbable if it were due only to chance.

## C. VIRULENCE AND VIABILITY OF STRAINS

The isolation of a number of new strains of TMV calls forth questions about their comparative virulence and viability. As previously mentioned, many of the new strains are more harmful to the host than the parental common TMV. Others are similarly, others less pathogenic, even to the point of giving no visible symptoms on two varieties of tobacco. Thus, virulence may be increased or decreased by mutation. However, viability, as frequently indicated by lesion size and measured by lesion productivity, seems to be invariably decreased by mutation. All natural strains were less productive than common TMV, and this difference is more pronounced for many chemically produced strains (Veldee and Fraenkel-Conrat, 1962). Thus, as one might expect, common TMV, the result of extensive selection, is probably the specimen most perfectly adapted to its medium within the range of all possible variations of that theme. All newly produced mutants may well have appeared spontaneously at some previous time and may have fallen by the wayside in the course of natural selection.

The first mutant to be elucidated by us in structural terms, #171, has illustrated one possible mechanism of biological evaluation of a chemical structure of the kind that would be required for such selection (Tsugita and Fraenkel-Conrat, 1960). Resistance to enzymes is probably a favorable attribute for viruses. The blocking of the N-terminus of TMV protein by an acetyl group, and the enzyme blocking action of proline near the C-terminus may contribute in this manner to the stability of the virus. The replacement of this proline by a leucine which was observed in strain #171 abolished its resistance toward the enzyme, carboxypeptidase A, and it appears very possible that the lesser biological viability of strain #171 can be largely attributed to this cause, and thus to the particular amino acid exchange in position 156 along the chain. Similarly simple explanations have not offered themselves in regard to other strains. It is important to remember, however, that we are able to analyze only the most viable mutants, and that we are thus forced to discriminate against all the more interesting protein alterations which render it nonfunctional.

## VII. The Coding Problem

One of the most important problems for an understanding of the mechanism of protein synthesis is the functional relationship between the genetic material and the protein. Considerable evidence has accumulated which indicates that the bulk of the genetic information in typical organisms is transmitted from parent to offspring through DNA. According to present information, the messages of DNA are transmitted by way of RNA (the so-called messenger RNA) and result in a specific protein structure, with soluble RNA probably acting as intermediary. In TMV as well as in other RNA viruses, the RNA was proven to be the direct bearer of the entire genetic information required to determine the amino acid sequence of its protein coat and its symptoms in the host. This was best illustrated by the mixed reconstitution experiments described previously. This concept was further supported by the study of chemical mutants, since one or very few changes in the nucleotide chain were able to effect changes in terms of biological symptoms and protein structure. Conclusive proof for the direct messenger activity of viral RNA has come from the recent experiments on protein synthesis in a cell-free *E. coli* system.

With a small virus like TMV, one should in principle be able to study both the structure of the RNA and that of the corresponding protein. If we assume that the linear nucleotide sequence in RNA is related to the amino acid sequence of the protein coat, and that the nature of the correspondence can be generalized, then TMV appears ideal for the study of the "coding problem" of protein synthesis, since we already possess experimental methods for the *in vitro* production of mutants and have knowledge of the complete amino acid sequence of this protein. Unfortunately, the chemical study of the RNA is only in its infancy, and the localization of mutagenic chemical events on the polynucleotide chain is not yet possible. Chemical studies of the protein of strains, on the other hand, has yielded definitive data, as reported in the preceding section. We shall now attempt to discuss and interpret these experimental results in the light of current theories and experimental facts pertaining to the problem of coding and information transfer.

In 1954 Gamow published the first theoretical approach to the problem how nucleotide sequences might code for amino acid sequences. The main assumptions of Gamow's scheme were the following. First, that a trinucleotide in DNA determined one amino acid. Second, that the trinucleotide fragment of DNA determining an amino acid and its neighbor "overlapped" by two-thirds of their length. Third, that the code was "degenerate," which means that more than one triplet may code for the same amino acid.

A triplet of nucleotides was assumed to code for one amino acid because pairs of the 4 nucleotides gave only 16 possibilities, while triplets gave 64 possibilities. The first number is not enough to code for 20 naturally occurring amino acids, while 64 possible combinations are more than necessary and permit degeneracy.

With the increase in our knowledge of amino acid sequences in many kinds of proteins, Gamow's overlapping scheme has been experimentally ruled out. Brenner (1957) has given a proof of the theoretical impossibility of any universal overlapping triplet code.

In 1957 Crick *et al.*, introduced a non-overlapping triplet code. The authors suggested that of the 64 triplets some "made sense," i.e., they represented an amino acid, while others did not. As an example, if A-G-U[3] stood for leucine and C-A-U[3] for valine, then G-U-C and U-C-A would be condemned to represent nothing ("nonsense"), so that . . . A-G-U-C-A-U . . . anywhere along the chain could be read only as Leu.Val. What made this theory most appealing was the fact that the 64 possible triplets were thus automatically reduced to 20 sense-making ones, in agreement with the number of common building blocks in proteins. The resulting code was obviously non-overlapping, and it was comma-free, since each nucleotide along the chain pertained to only one sense-making and thus coding triplet. It should be noted that in a commaless code triplets composed of the same nucleotide (i.e., A-A-A) are excluded, since a sequence such as $-A^1-A^2-A^3-A^4-A^5-$ does not give a unique sense; for it can be read starting at position 1, 2, or 3.

More recently, Crick *et al.* (1961) presented the first experimental evidence, based on genetic studies of T4 phage and acridine mutants derived from it, that the code for the DNA of bacterial genes was actually triplet in nature. They also arrived at the conclusion that the reading of the code must be from one end, and thus does not require nonsense triplets as commas. Their new data seemed to rule out the existence of meaningless triplets. Only between cistrons was there evidence for a break in the continuity of decipherable nucleotide sequences.

In addition to the theoretical and genetic approaches to the coding problem, an analytical approach was also attempted. Thus, Yčas proposed a code based on correlations between published analyses of the nucleotide and amino acid compositions of simple viruses. Woese (1961a,b) followed this path further, additionally using amino acid exchange data from several kinds of proteins in related species.

A great advance in our understanding of the coding problem has recently come about when Nirenberg and Matthaei reported on their use of an excellent system to attack this problem in direct and experimental manner (1961). We have previously described the cell-free protein synthesizing system of *E. coli* (containing S-RNA, ribosomes, an ATP-generating system, and $C^{14}$-amino acids) which these authors reported to be greatly stimulated by the addition of any RNA, and particularly by TMV-RNA. Surprisingly, polyuridylic acid (poly-U) was found to be particularly effective in this regard, and Nirenberg and Matthaei were soon able to demonstrate that this additive led to the exclusive synthesis of polyphenylalanine.

This finding attracted the deserved attention, and other laboratories quickly initiated research projects along similar lines and reported preliminary results. Thus Lengyel *et al.* (1961), using essentially the same system, confirmed that poly-U led to the synthesis of polyphenylalanine. These workers obtained a molar ratio of 3.25 between the amount of nucleotide used and the newly incorporated amino acid, compared to 2.4 reported by Nirenberg and Matthaei (1961). This figure may represent a first approximation by biochemical methods to the coding ratio, and it is heartening that it is so close to the triplet recently postulated by Crick *et al.* (1961) on genetic grounds.

---

[3] A represents a nucleotide residue, the base of which is adenine; G represents guanine; U, uracil; and C, cystosine.

Further amino acid incorporation studies in the laboratories of both Nirenberg and Ochoa have centered on the use of various nucleotide copolymers of different composition (Martin *et al.*, 1962; Speyer *et al.*, 1962a,b; Lengyel *et al.*, 1962). These favored in more or less specific manner the incorporation of other amino acids, but, as one might expect, these random copolymers were quantitatively much less effective than poly-U. All polymers that have been reported on contained uridylic acid and their effectiveness in regard to the incorporation of various amino acids was generally expressed in terms of the ratio to phenylalanine incorporation. Although the use of random copolymers did not enable the authors to determine the sequential structure of any coding unit, except for poly-U, nor whether the unit was definitely a triplet, they were able to suggest probable coding combinations for 17 out of the 20 naturally occurring amino acids, with only a few instances of disagreement in the results of the two laboratories participating in the race. In addition, Nirenberg *et al.* (1962) have shown that phenylalanine linked to soluble RNA was an intermediate in the process, using polyuridylic acid.

The finding that poly-U carries the message for phenylalanine polymerization is in disagreement with Crick's "commaless triplet code" as described. As mentioned before, these facts and their own data have led Crick *et al.* (1961) to modify their concepts and to abandon the postulation of nonsense triplets. It thus becomes ever more likely that the code is a degenerate one.

Two years preceding these exciting demonstrations, we reported the amino acid exchanges in a chemically evoked mutant of TMV (Tsugita and Fraenkel-Conrat, 1960), and Wittmann (1959b) observed no amino acid exchanges in another chemical mutant. Since then both groups have concentrated their efforts on analyzing many mutant proteins, and these observations have supplied important experimental data pertaining to the coding problem.

Wittmann analyzed 120 nitrous acid mutants (1961) and we analyzed 56 chemically induced mutants (Tsugita and Fraenkel-Conrat, 1960, 1962; Tsugita, 1962a,b). In both laboratories at least half of the observed mutants showed no amino acid exchange. We have previously discussed the grounds for accepting this negative finding as valid. This finding means that a good part of the RNA does not find expression in the composition of the virus coat protein. Thus, coding schemes based on the direct comparison between the composition of the RNA and the protein of the viruses, such as those proposed by Yčas and Woese, seem to have lost any basis in reality. The finding that not all RNA alterations lead to coat protein alterations is in line with expectation, if one considers the numbers of residues involved. Assuming a triplet code, the 158 amino acids would be expected to be governed by only 474 nucleotides, which corresponds to only 7.4% of the nucleotides in TMV-RNA. Most of the RNA would thus be assumed to carry information not pertaining to the coat protein, which alone has been analyzed. Actually one might expect that only about 1 out of 10 mutants should show protein changes, but it is quite possible that a high proportion of the particles altered in segments pertaining to properties other than protein coat structure are not viable.

The chemical mutant (#171) first analyzed by us (Tsugita and Fraenkel-Conrat, 1960) had 3 amino acid exchanges and the location of these changes was scattered throughout the sequence. This observation, as well as the detection of mutants with single exchanges, disproves any overlapping code.

Investigations of the deamination reaction with DNA have led Vielmetter and Schuster (1960) to the conclusion that only the deamination of cytosine and adenine was mutagenic. The resulting uracil and hypoxanthine would on chemical grounds be expected to pair preferentially as if they were thymine and guanine, respectively. In contrast, the deamination product of guanine, xanthine, still resembles guanine. Based on such consideration and an assumed triplet code, Gierer (1962) arranged all possible permutations of the 64 triplets in a scheme governed by the result of their deamination (C → U, A → (hypoxanthine) → G) in the form of 8 octets as shown in the diagram.

$$\begin{tikzcd}
& \mathrm{CAC} \arrow[dl] \arrow[d] \arrow[dr] & \\
\mathrm{UAC} \arrow[d] \arrow[dr] & \mathrm{CGC} \arrow[dl] \arrow[dr] & \mathrm{CAU} \arrow[dl] \arrow[d] \\
\mathrm{UGC} \arrow[dr] & \mathrm{UAU} \arrow[d] & \mathrm{CGU} \arrow[dl] \\
& \mathrm{UGU} &
\end{tikzcd}$$

It is evident that this scheme sets strict limitations on the possible results of deamination. Thus, certain amino acids might be replaced by three others but none by more. Any amino acid ever observed to result from an exchange could only be exchanged to two or one other. Finally, no amino acid could result from more than three others, and those in this latter group would never be replaced. Wittmann has relied on this scheme in the interpretations of his results, but he finds some exchanges that do not fit, and we have observed others.

As previously discussed, we believe that in RNA, as contrasted to DNA, the deamination of cytosine is the greatly predominating mutagenic event. Only much more rarely might the deamination of A cause an A → G shift, and other base exchanges may arise with even lower frequency during replication of the deaminated as well as the brominated or methylated RNA, and attributable to random mistakes in replication in a manner similar to natural mutation. Thus, the octet scheme appears too rigid to allow for all possible events, and not truly expressive of the predominant role of the C → U change in RNA deamination.

Focusing our attention on this predominant result of deamination, we have arranged our observed amino acid exchanges schematically in groups headed by triplets of decreasing cytosine content, allocating amino acids to triplets on the basis of the recent coding findings discussed above (see Fig. 8). The implication in such an evaluation of our exchange data is that they are the consequence of single step mutations, which is probably the case for most single exchanges, but can in no individual case be stated with assurance.

Further, one must consider whether the coding triplets as observed with *E. coli* would be of the same "language" as those in tobacco plant cells, since there exists the possibility that different codes might pertain in different organisms. This question appears to be answered by the results reported in a preceding section where it was shown that TMV-RNA acted as messenger for the production of material resembling the TMV protein by the same *E. coli* system, which has been used to obtain the coding data with synthetic polynucleotides. Thus, the coding language, whether it is universal or not, appears to be very similar for *E. coli* and tobacco, and the correlations shown in Fig. 8 are thus justified.

The data presented make it appear probable that proline, threonine and aspartic acid or its amide are coded as $C_2U$; leucine, serine, and alanine as $CU_2$; and phenylalanine as $U_3$. These hypotheses agree in part with the results obtained by one and/or the other group studying the amino acid incorporation in the *E. coli* system. Cases which do not agree with these data (e.g., Tyr → Phe, with tyrosine being coded by an

| | | |
|---|---|---|
| $C_2U$ | $CU_2$ | $U_3$ |
| $C_2X$ | CXU | $XU_2$ |
| | $CX_2$ | $X_2U$ |

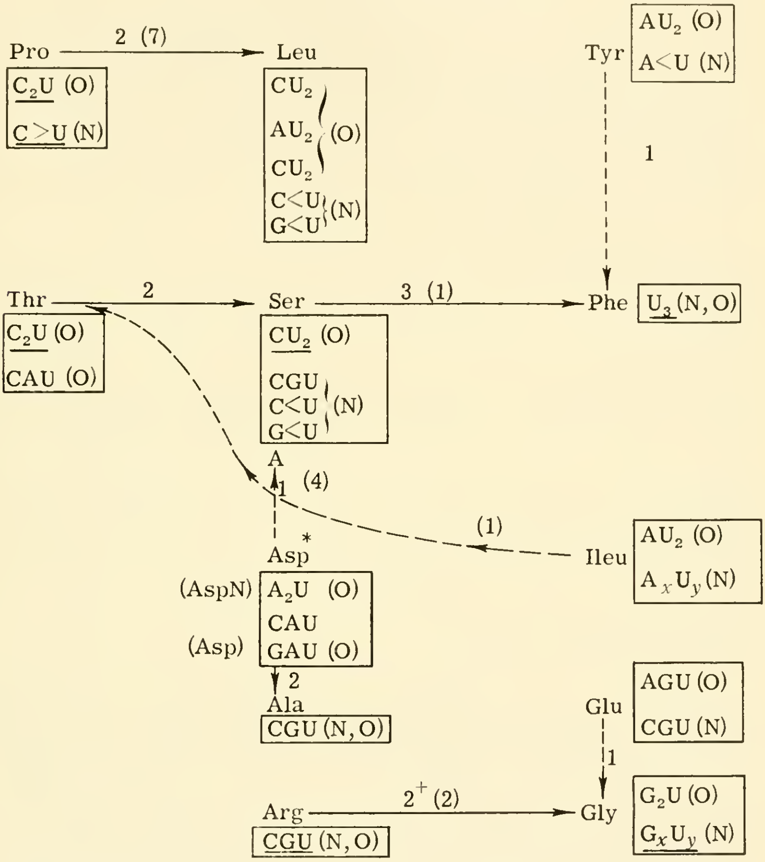

Pro
2 (7)
Leu
$C_2U$ (O)
C>U (N)
$CU_2$
$AU_2$ (O)
$CU_2$
C<U
G<U (N)
Tyr
$AU_2$ (O)
A<U (N)
1
Thr
2
Ser
3 (1)
Phe
$U_3$(N,O)
$C_2U$ (O)
CAU (O)
$CU_2$ (O)
CGU
C<U (N)
G<U
A
1 (4)
(1)
*
Asp
Ileu
$AU_2$ (O)
$A_xU_y$ (N)
(AspN)
$A_2U$ (O)
CAU
(Asp)
GAU (O)
2
Ala
CGU (N,O)
Glu
AGU (O)
CGU (N)
1
Arg
$2^+$ (2)
Gly
CGU (N,O)
$G_2U$ (O)
$G_xU_y$ (N)

FIG. 8. Correlation of observed amino acid replacements with coding symbols. All amino acid exchanges that have been observed in this laboratory are shown, with their frequency of occurrence. Numbers in parenthesis indicate exchanges resulting from reactions other than deamination, and thus unrelated to the arrangement scheme of this presentation which is based on the decreasing cytosine content from left to right. Solid lines and underlined code symbols indicate agreement of amino acid exchange data with C → U transformations. "N" and "O" after the code symbols refers to the origin of the respective identification, "N" standing for Nirenberg and Matthaei (1961), Nirenberg *et al.* (1962), and Martin *et al.* (1962), and "O" standing for Lengyel *et al.*, 1961, 1962, and Speyer *et al.* (1962). The symbol $2^+$ in the bottom line refers to the fact that a progeny isolate of this strain showed three arginines to be replaced by glycines, rather than two.

A-containing nucleotide, according to both groups) must be attributed to mutational events other than C → U and A → G which, as stated, are definitely possible. Alternately, the degeneracy of the code might supply an explanation. It is encouraging, however, that at least 55% of the amino acid exchanges observed by us can be attributed to cytosine deaminations on the basis of present coding data. About 15% of the changes suggest an A → G exchange, but an equal fraction is indicative of C → G and A → U exchanges. On the other hand, the fact that the allotment of code symbols to amino acids has been influenced by exchange data obtained with nitrous acid mutants naturally renders the "agreement" less surprising. It also suggests that the limits in the useful exploitation of random nucleotide copolymers for coding information have now been reached, if not overstepped.

To conclude this short survey of the present concepts concerning the code that relates RNA to protein structure, one is now in the fortunate position that much experimental data has recently become available which brings this problem from the realm of speculation to that of factual information. It appears that three or possibly a slightly greater number of nucleotides code for one amino acid; it appears probable that several oligonucleotide units may represent the same amino acid. It is certain that a few uridylic acids represent one phenylalanine, and the approximate nucleotide composition of the words spelling 16 other amino acids are known with varying certainty. The code seems to be quite similar for very different organisms. It is probably deciphered starting from an end of the nucleotide chain and thus requires no markers (commas) between "words."

The main projects for the immediate future are (*1*) the preparation, by chemical or biosynthetic means, of polymers of known repeating nucleotide sequences and their systematic testing for messenger activity, (*2*) the establishment of the nature of many decidedly single-hit, one-step mutations in terms of resulting amino acid exchanges, and (*3*) the development of methods to determine the nucleotide exchange produced by the mutation, be it in the DNA or in the messenger RNA.

It would appear that within a few years it may become possible to produce a known modification in one of the simpler DNA or RNA genomes and find it to result in a predicted protein alteration. The natural mode of transfer of information from the genetic DNA to the messenger RNA is now reasonably well understood. There remains only one major obstacle to overcome before we can fruitfully exploit our new knowledge toward the manipulation of even the simplest genetic material, which is that of the small viruses. The remaining hurdle is the lacking methodology for the sequential analysis of long polynucleotides, a feat which appears to be the prerequisite for the exact superposition of the genetic map on the polynucleotide chain, the crowning achievement in molecular genetics.

## REFERENCES

Aach, H. G. (1958). *Z. Naturforsch.* **13b**, 426.

Anderer, F. A. (1959). *Z. Naturforsch.* **14b**, 642.

Anderer, F. A. (1962). *5th Intern. Congr. Biochem., Moscow, 1961.*

Anderer, F. A., Weber, E., and Uhlig, H. (1960a). *Z. Naturforsch.* **15b**, 79.

Anderer, F. A., Uhlig, H., Weber, E., and Schramm, G. (1960b). *Nature* **196**, 922.

Ansevin, A. T., and Lauffer, M. A. (1959). *Nature* **183**, 1601.

Bawden, F. C., and Pirie, N. W. (1937). *Proc. Roy. Soc.* **B123**, 274.

Beijerinck, M. W. (1899). *Zentr. Bakteriol. Parasitenk. Abt. II* **5**, 27.

Black, F. L., and Knight, C. A. (1953). *J. Biol. Chem.* **202**, 51.

Blout, E. R., DeLoze, C., Bloom, S. M., and Fasman, G. W. (1960). *J. Am. Chem. Soc.* **82**, 3787.

Boedtker, H. (1959). *Biochim. et Biophys. Acta* **32**, 519.

Bonhoeffer, F., and Schachman, H. K. (1960). *Biochem. Biophys. Research Communs.* **2**, 366.

Braunitzer, G. (1956). *Biochim. et Biophys. Acta* **19**, 574.

Brenner, S. (1957). *Proc. Natl. Acad. Sci. U.S.* **43**, 416.

Caspar, D. L. D. (1956). *Nature* **177**, 928.

Colter, J. S. (1958). *Progr. in Med. Virol.* **1**, 1.

Crick, F. H. C., Griffith, J. S., and Orgel, L. E. (1957). *Proc. Natl. Acad. Sci. U.S.* **43**, 687.

Crick, F. H. C., Barnett, L., Brenner, S., and Watts-Tobin, R. J. (1961). *Nature* **192**, 1227.

d'Hérelle, F. (1917). *Compt. rend. acad. sci.* **165**, 373.

Doty, P., Boedtker, H., Fresco, J. R., Haselkorn, R., and Litt, M. (1959). *Proc. Natl. Acad. Sci. U.S.* **45**, 482.

Engler, R., and Schramm, G. (1959). *Nature* **183**, 1277.

Fraenkel-Conrat, H. (1955a). *J. Biol. Chem.* **217**, 375.

Fraenkel-Conrat, H. (1955b). *Biochim. et Biophys. Acta* **15**, 307.

Fraenkel-Conrat, H. (1956). *J. Am. Chem. Soc.* **78**, 882.

Fraenkel-Conrat, H. (1957). *Virology* **4**, 1.

Fraenkel-Conrat, H. (1959). *In* "Sulfur in Proteins" (R. Benesch *et al.,* ed.), p. 339. Academic Press, New York.

Fraenkel-Conrat, H. (1961). *Biochim. et Biophys. Acta* **49**, 169.

Fraenkel-Conrat, H. (1962). *In* "Survey of Biological Progress," (B. Glass, ed.), Vol. IV, p. 59. Academic Press, New York.

Fraenkel-Conrat, H., and Narita, K. (1958). *Symposium on Protein Structure, Paris, 1957,* p. 249.

Fraenkel-Conrat, H., and Ramachandran, L. K. (1959). *Advances in Protein Chemistry* **14**, 175.

Fraenkel-Conrat, H., and Singer, B. (1957). *Biochim. et Biophys. Acta* **24**, 540.

Fraenkel-Conrat, H., and Singer, B. (1959). *Biochim. et Biophys. Acta* **33**, 370.

Fraenkel-Conrat, H., and Singer, B. (1962). *Biochemistry* **1**, 120.

Fraenkel-Conrat, H., and Williams, R. C. (1955). *Proc. Natl. Acad. Sci. U.S.* **41**, 690.

Fraenkel-Conrat, H., Singer, B., and Williams, R. C. (1957). *Biochim. et Biophys. Acta* **25**, 87.

Fraenkel-Conrat, H., Singer, B., and Veldee, S. (1959). *Biochim. et Biophys. Acta* **29**, 639.

Fraenkel-Conrat, H., Singer, B., and Tsugita, A. (1961). *Virology* **14**, 59.

Franklin, R. E., and Holmes, K. C. (1958). *Acta Cryst.* **11**, 213.

Franklin, R. E., and Klug, A. (1955). *Acta Cryst.* **8**, 777.

Franklin, R. E., and Klug, A. (1956). *Biochim. et Biophys. Acta* **19**, 403.

Franklin, R. E., Klug, A., and Holmes, K. C. (1957). "The Nature of Viruses" (G. E. W. Wolstenholme and E. C. P. Millar, eds.), p. 39. Churchill, London.

Frisch-Niggemeyer, W. (1956). *Nature* **178**, 307.

Gamow, G. (1954). *Nature* **173**, 318.

Gierer, A. (1957). *Nature* **179**, 1297.

Gierer, A. (1958a). *Z. Naturforsch.* **13b**, 477.

Gierer, A. (1958b). *Z. Naturforsch.* **13b**, 485.

Gierer, A. (1960). *Progr. in Biophys. and Biophys.* **10**, 299.

Gierer, A. (1962). *5th Intern. Congr. Biochem., Moscow, 1961.*

Gierer, A., and Mundry, K. W. (1958). *Nature* **182**, 1457.

Gierer, A., and Schramm, G. (1956). *Z. Naturforsch.* **11b**, 138 (1956).

Ginoza, W. (1958). *Nature* **181**, 958.

Gish, D. T. (1959). *Biochim. et Biophys. Acta* **35**, 557.

Gish, D. T. (1960). *J. Am. Chem. Soc.* **82**, 6329.

Gish, D. T. (1961). *J. Am. Chem. Soc.* **83**, 3303.

Gish, D. T., Ramachandran, L. K., and Stanley, W. M. (1958). *Arch. Biochem. Biophys.* **78**, 433.

Gordon, M. P., and Staehelin, M. (1958). *J. Am. Chem. Soc.* **80**, 2340.

Harris, J. I., and Knight, C. A. (1952). *Nature* **170**, 613.

Harris, J. I., and Knight, C. A. (1955). *J. Biol. Chem.* **214**, 215.

Hart, R. G. (1958). *Biochem. et Biophys. Acta* **28**, 457.

Haschemeyer, R., Singer, B., and Fraenkel-Conrat, H. (1959). *Proc. Natl. Acad. Sci. U.S.* **45**, 313.

Haselkorn, R., and Doty, P. (1961). *J. Biol. Chem.* **236**, 2738.

Hershey, A. D., and Chase, M. (1952). *J. Gen. Physiol.* **36**, 39.

Iwanowski, D. (1892). *St. Petersburg Acad. Imp. Sci. Bull.* **35**, 67. (Reprinted in English (1942): *Phytopathol. Classics* **7**, 27.)

Jeener, R., and Rosseels, J. (1953). *Biochim. et Biophys. Acta* **11**, 438.

Klug, A., and Caspar, D. L. D. (1960). *Advances in Virus Research* **7**, 225.

Knight, C. A. (1947). *J. Biol. Chem.* **171**, 297.

Knight, C. A. (1952). *J. Biol. Chem.* **197**, 241.

Knight, C. A. (1955). *Virology* **1**, 261.

Knight, C. A. (1959). *In* "The Viruses" (F. M. Burnet and W. M. Stanley, eds.), Vol. 2, p. 127. Academic Press, New York.

Knight, C. A., and Stanley, W. M. (1941). *J. Biol. Chem.* **141**, 39.

Knight, C. A., Silva, D. M., Dahl, D., and Tsugita, A. (1962). *Virology* **16**, 236.

Kramer, G., and Wittmann, H. G. (1958). *Z. Naturforsch.* **13b**, 30.

Lengyel, P., Speyer, J. F., and Ochoa, S. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 1936.

Lengyel, P., Speyer, J. F., Basilio, C., and Ochoa, S. (1962). *Proc. Natl. Acad. Sci. U.S.* **48**, 282.

Littlefield, J. W., and Dunn, D. B. (1958). *Biochem. J.* **70**, 642.

Mandel, H. G., Markham, R., and Matthews, R. E. F. (1957). *Biochim. et Biophys. Acta* **24**, 205.

Markham, R., and Smith, J. D. (1950). *Biochem. J.* **46**, 513.

Martin, R. G., Matthaei, J. H., Jones, O. W., and Nirenberg, M. W. (1962). *Biochem. Biophys. Research Communs.* **6**, 410.

Matthews, R. E. F., and Smith, J. D. (1955). *Advances in Virus Research* **3**, 49.

Mundry, K. W. (1959). *Virology* **9**, 722.

Narita, K. (1958). *Biochim. et Biophys. Acta* **28**, 184.

Newmark, P., and Meyers, R. W. (1957). *Federation Proc.* **16**, 226.

Nirenberg, M. W., and Matthaei, J. H. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 1588.

Nirenberg, M. W., and Matthaei, J. H., and Jones, O. W. (1962). *Proc. Natl. Acad. Sci. U.S.* **48**, 104.

Niu, C. I., and Fraenkel-Conrat, H. (1955a). *Biochim. et Biophys. Acta* **16**, 597.

Niu, C. I., and Fraenkel-Conrat, H. (1955b). *Arch. Biochem. Biophys.* **59**, 538.

Nixon, H. L., and Woods, R. D. (1960). *Virology* **10**, 157.

Ramachandran, L. K., and Gish, D. T. (1959). *J. Am. Chem. Soc.* **81**, 884.

Rushizky, G. W., and Knight, C. A. (1960). *Virology* **11**, 236; *Proc. Natl. Acad. Sci. U.S.* **46**, 945.

Schaffer, F. L., and Schwerdt, C. E. (1955). *Proc. Natl. Acad. Sci. U.S.* **41**, 1020.

Schramm, G., and Zillig, W. (1955). *Z. Naturforsch.* **9b**, 494.

Schramm, G., Schumacher, G., and Zillig, W. (1955). *Z. Naturforsch.* **10b**, 481.

Schuster, H. (1960). *In* "The Nucleic Acids" (E. Chargaff and J. N. Davidson, eds.), Vol. 3, p. 245. Academic Press, New York.

Schuster, H. (1961). *J. Mol. Biol.* **3**, 447.

Schuster, H., and Schramm, G. (1958). *Z. Naturforsch.* **13b**, 697.

Siegel, A. (1960). *Virology* **11**, 156.

Siegel, A., Ginoza, W., and Wildman, S. G. (1957). *Virology* **3**, 554.

Singer, B., and Fraenkel-Conrat, H. (1961). *Virology* **14**, 59.

Speyer, J. F., Lengyel, P., Basilio, C., and Ochoa, S. (1962a). *Proc. Natl. Acad. Sci. U.S.* **48**, 63.

Speyer, J. F., Lengyel, P., Basilio, C., and Ochoa, S. (1962b). *Proc. Natl. Acad. Sci. U.S.* **48**, 441.

Staehelin, M. (1958). *Biochim. et Biophys. Acta* **29**, 410.

Staehelin, M. (1959a). *Biochim. et Biophys. Acta* **31**, 448.

Staehelin, M. (1959b). *Experientia* **15**, 413.

Staehelin, M., Peterson, E. A., and Sober, H. A. (1959). *Arch. Biochem. Biophys.* **85**, 289.

Stanley, W. M. (1935). *Science* **81**, 644.

Stanley, W. M. (1936). *Science* **83**, 626.

Stanley, W. M. (1938). *In* "Handbuch der Virusforschung" (R. Doerr and C. Hallauer, eds.), Vol. IV, p. 447. Springer, Berlin.

Sugiyama, T., and Fraenkel-Conrat, H. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 1393.

Sugiyama, T., and Fraenkel-Conrat, H. (1962). In preparation.

Sugiyama, T., Singer, B., and Fraenkel-Conrat, H. (1962). In preparation.

Takahashi, W. N. (1959). *Virology* **9**, 437.

Takahashi, W. N., and Ishii, M. (1952). *Nature* **169**, 419.

Tsugita, A. (1960). *Biochim. et Biophys. Acta* **38**, 145.

Tsugita, A. (1961). *Protein Nucleic Acid, Enzyme (Tokyo)* **6**, 385.

Tsugita, A. (1962a). *J. Mol. Biol.*, in press.

Tsugita, A. (1962b). *J. Mol. Biol.*, in press.

Tsugita, A., and Fraenkel-Conrat, H. (1960). *Proc. Natl. Acad. Sci. U.S.* **46**, 636.

Tsugita, A., and Fraenkel-Conrat, H. (1962). *J. Mol. Biol.* **4**, 73.

Tsugita, A., Gish, D. T., Young, J., Fraenkel-Conrat, H., Knight, C. A., and Stanley, W. M. (1960). *Proc. Natl. Acad. Sci. U.S.* **46**, 1463.

Tsugita, A., Funatsu, G., and Young, J. (1962a). In preparation.

Tsugita, A., Fraenkel-Conrat, H., Nirenberg, M. W., and Matthaei, J. H. (1962b). *Proc. Natl. Acad. Sci. U.S.* **48**, 846.

Twort, F. W. (1915). Lancet **ii**, 1241.

Veldee, S., and Fraenkel-Conrat, H. (1962). *Virology,* in press.

Vielmetter, W., and Schuster, H. (1960). *Z. Naturforsch.* **15b**, 304.

Whitfeld, P. (1962). In preparation.

White, F. H., Jr., and Anfinsen, C. B. (1959). *Ann. N.Y. Acad. Sci.* **81**, 515.

Wildman, S. G. (1959). *In* "The Viruses" (F. M. Burnet and W. M. Stanley, eds.), Vol. 2, p. 1. Academic Press, New York.

Wittmann, H. G. (1959a). *Z. Vererbungslehre* **90**, 463.

Wittmann, H. G. (1959b). *Experientia* **15**, 174.

Wittmann, H. G. (1960a). *Virology* **11**, 505.

Wittmann, H. G. (1960b). *Virology* **12**, 613.

Wittmann, H. G. (1960c). *Virology* **12**, 609.

Wittmann, H. G. (1961). *Naturwiss.* **48**, 729.

Wittmann, H. G., and Braunitzer, G. (1959). *Virology* **9**, 762.

Woese, C. R. (1961a). *Nature* **190**, 697.

Woese, C. R. (1961b). *Biochem. Biophys. Research Communs.* **5**, 88.

Yčas, M. (1960). *Nature* **188**, 209.

Zech, H. (1961). *Z. Naturforsch.* **16b**, 520.

# AUTHOR INDEX

Numbers in italics indicate the pages on which the references are listed at the end of the article.

A

Aach, H. G., 493, *517*
Abbo, F. E., 155, 159, *200*
Abelson, P. H., 178, *204*
Abrahamov, A., 461, *470*
Acs, G., 356, 357, 358, *402, 404*
Adelberg, E. A., 146, *151,* 218, *265*
Adler, J., 13, 14, 23, 24, 29, 40, 41, 42, 43, 44, 45, 46, 48, *60, 61, 63*
Ager, J. A. M. 454, *469*
Aksoy, M., 436, *469*
Alberts, B. M., 247, *266*
Albertsson, P.-A., 310, *350*
Alexander, P., 232, 233, 237, 238, *264, 268*
Alfert, M., 92, *108,* 157, 190, 194, *200, 201*
Allen, E. H., 355, 356, 358, 361, 362, 364, 365, 366, 367, 375, 399, *400, 401, 403*
Allen, D. W., 408, 443, 459, 467, *469, 470*
Allfrey, V. G., 395, 397, *400, 401, 404*
Allison, A. C., 421, 463, *469, 471*
Alpen, E. L., 466, 467, *470*
Alper, H., 243, *264*
Ames, B. N., 119, *148*
Anderegg, J. W., 13, *61,* 173, *200*
Anderer, F. A., 482, 484, 485, 486, *517*
Anderson, B., 13, 17, *62*
Anderson, E. S., 120, *148*
Anderson, R. S., 242, *266*
Anderson, T. F., 292, *351*
Andrew, L. E., 218, *264*
Anfinsen, C. B., 354, 373, *400, 404,* 482, *520*
Ansevin, A. T., 480, *517*
Apgar, J., 361, *401*
Aposhian, H. V., 55, 56, *60*
Arber, W., 133, *148*
Arient, M., 165, 166, *204*
Arlinghaus, R., 370, 379, *403*
Armstrong, J. A., 120, *148*
Arnstein, H., *393*
Aronson, A. I., 179, 187, *200,* 298, 305, 315, 329, *350, 351*
Askonas, B., 373, *400*
Astrachan, L., 179, *200,* 273, 275, 276, 278, 279, *286, 289,* 336, 337, *350, 352,* 386, *400*
Attardi, G., 384, 386, *402*
Awater, J., 437, 438, 439, 440, 441, 450, 451, *470*
Atwood, K. C., 245, *268*
Aub, J. C., 191, *200*
Auerbach, C., 210, 234, 235, 240, *264, 268*
Aurisicchio, S., 119, *148*
Austin, J. H., 416, *471*
Avery, O. T., 72, *108*

B

Baddiley, J., 166, *200*
Baglioni, C., 410, 411, 412, 413, 416, 423, 424, 425, 427, 434, 438, 439, 440, 443, 446, 448, 449, 454, 463, *470, 474*
Bailey, I. S., 451, *470*
Baker, W. K., 243, *264*
Baldesten, A., 3, *63*
Baldwin, R. L., 50, *60*
Balis, M .E., 185, *200*
Balog, J. E., 408, 441, 443, 453, *469, 472*
Balog-Shelton, J., 416, 417, *474*
Barclay, R. K., 125, *149*
Barker, T., 176, *202*
Barlow, J. L., 119, 120, *150*
Barner, H. D., 9, *60, 61,* 143, *148,* 156, 158, 170, 174, 194, *200, 201*
Barnes, D. W. H., 466, *470*
Barnes, H. B., 275, *287*
Barnett, L., 225, 248, 249, 250, 253, 255, 260, *265,* 285–286, *287,* 377, 390, *401,* 435, 436, *470,* 511, 512, *517*
Barnicot, N. A., 445, *474*
Barnum, C. P., 166, 191, *200*
Baron, F., 227, *264*
Baron, L. S., 147, *148, 149, 150*
Bartuska, D., *204*
Basch, R. S., 165, *202*
Basilio, C., 282, 283, *288,* 390, 391, 392,

*400, 404,* 434, 435, 436, *474,* 512, 515, *518, 519*
Bass, L. W., 227, *265*
Bauch, R., 225, *264*
Bautz, E., 222, 228, 229, 230, 234, 237, 238, 239, 249, 250, 251, 253, 255, 261, 263, *264, 266,* 279, *286,* 381, 389, *400*
Bawden, F. C., 478, *517*
Bayley, C. R., 230, *264*
Bearn, A. G., 418, 420, *472*
Beaven, G. H., 408, 416, 418, 428, 455, 459, 461, *470, 471, 474*
Beer, M., 126, *148*
Beers, R. F., 225, *264, 269*
Beiboer, J. L., 447, *473*
Beijerinck, M. W., 477, *517*
Bell, D., 13, *62,* 171, 172, *202, 205*
Bello, L. J., 16, *60*
Belser, N. D., 247, *269*
Bendet, I., 114, 126, *148*
Bendich, A., 107, *108*
Benesch, R., 454, *470*
Benesch, R. E., 454, *470*
Ben-Ishai, R., 384, *404*
Bentzon, M. W., 155, 195, *205*
Benzer, S., 121, *150,* 211, 212, 214, 219, 222, 225, 226, 238, 248, 249, 250, 252, 253, 255, *264, 265,* 361, *400,* 424, 426, *470*
Berends, W., 245, *265*
Berg, P., 17, *60,* 280, 283, 284, *287, 289,* 354, 355, 356, 358, 359, 360, 361, 380, 384, 388, 389, 394, *400, 401, 403, 404*
Bergmann, F., 356, 358, 359, 361, *400,* 403
Bergren, W. R., 422, *472*
Bernal, J. D., 114, *148*
Bernardi, G., 145, *148*
Berns, K. I., 128, 129, 130, 138, *148, 151*
Bertani, E., 3, 4, 5, 9, 50, *60*
Bessman, M. J., 13, 14, *15,* 16, 20, 21, 23, 24, 26, 27, 29, 31, 32, 33, 46, 47, 48, 50, *60, 61, 63,* 160, *200, 203,* 273, 280, *286, 287*
Betke, K., 434, 459, 461, 462, *470, 472*
Beukers, R., 245, *265*
Bevelander, G., 165, *202*
Bianchi, P. A., 13, 14, *61,* 171, 172, 174, *200*
Bianco, I., 418, 447, 449, 454, *474*
Billen, D., 173, 174, 176, 178, 179, 181, 190, 191, *200, 201*
Binkley, S. B., 170, *201*
Binnie, B., 186, *202*
Biringer, R., 280, *287*
Birnie, G. D., 7, *61*
Bishop, C. N., 165, 166, *201*
Bishop, J., 363, 364, 365, 366, 367, 371, 372, 374, 375, 379, 386, 394, 398, 399, *400, 403*
Black, F. L., 493, *517*
Blakely, R. L., 7, *61*
Blakeslee, A. F., 248, *265*
Bloch, D. P., 92, 96, *109,* 157, 188, *200*
Bloemendal, H., 368, 369, *400, 401*
Bloom, S. M., 482, *517*
Bloom, W., 459, *473*
Blout, E. R., 482, *517*
Blumson, N. L., 166, *200*
Boedtker, H., 487, 488, *517*
Boehner, L., 136, 137, *149*
Boezi, J. A., 277, *287,* 343, *350*
Boivin, A., 72, *109*
Bojarski, R. B., 14, *61, 62,* 171, 172, 174, *200, 202*
Bolle, A., 179, 180, 186, 187, *203*
Bollum, F. J., 13, 50, 51, 52, 53, 54, *61,* 171, 172, 173, 190, *200, 204*
Bolton, E. T., 178, *204,* 277, *287,* 294, 298, 301, 302, 303, 305, 306, 309, 313, 315, 333, 336, *350, 352*
Bolton, T. T., 305, *350*
Boman, H., 356, 358, *402*
Bond, V. P., 169, *201,* 396, *401*
Bonhoeffer, F., 487, *517*
Bonner, D. M., *288*
Bonner, J., 186, *200,* 279, 280, *287,* 292, 294, 301, *352,* 395, 396, 397, *401, 403*
Borek, E., 10, *63*
Borka, T., 157, *203*
Borsook, H., 299, 301, *350, 351,* 353, 354, 374, *401*
Bosch, L., 368, *400, 401*
Boulanger, P., 277, *287*
Boulard, C., *446, 470*
Bovard, F., 358, *403*
Bowman, H. G., 24, *61*
Boyce, R., 184, *205*
Boyer, S. H., 441, 442, 443, 446, 454, *474*
Boyo, A. E., 454, *471*

Brachet, J., 272, *287*, 299, *350*, 395, 396, *401*
Bradley, D. F., 225, *265*
Bradley, T. B., Jr., 462, 463, *470*
Brakke, M. K., 309, *350*
Brammer, K. W., 230, *264*
Brandt, P. W., 97, 104, 105, *110*
Braun, H., 461, 462, *472*
Braun, W., 196, *201*
Braunitzer, G., 413, 414, 415, 416, 417, *470*, *473*, 484, 485, 501, *517*, *520*
Brawner, J. N., III, 462, 463, *470*
Brenner, S., 180, 186, *200*, *205*, 222, 225, 248, 249, 250, 253, 255, 260, *265*, *268*, 274, 285–286, *287*, 310, 344, *350*, 377, 383, 390, *401*, 435, 436, *470*, 510, 511, 512, *517*
Bresler, A., 280, *287*, 347, 348, *351*, 382, *402*
Bridges, C. B., 448, *470*
Bridges, M. T., 444, 448, *471*
Britten, R. J., 178, 179, *200*, *204*, 277, *287*, *288*, 296, 297, 299, 301, 305, 309–310, 312, 313, 315, 322, 323, 333, 344, *350*, *351*, *352*
Brodie, H. R., 459, *470*
Brody, S., 185, *200*
Bromberg, Y. M., 461, *470*
Brookes, J., 237, *265*
Brown, A., 358, *401*
Brown, B. J., 156, 174, 175, 184, *202*, 375, *401*
Brown, D. M., 227, 228, *264*, *265*
Brown, G., 358, *401*
Brown, G. B., 243, *269*
Brown, G. L., 125, *149*
Brownell, L. W., 165, 166, *205*
Bruce, V., 131, 132, *151*
Buchanan, J. M., 5, 50, 55, 57, *62*, *63*
Bucher, N. R. L., 191, 197, *200*
Buchwald, M., 313, *351*
Buettner-Janusch, T., 165, 166, *204*
Buettner-Janusch, V., 12, *63*, 165, *166*, *204*
Bulivant, S., 107, *109*
Bulman, H. N., 196, *201*
Burckett, L., 462, 463, *472*
Burgi, E., 122, 123, 125, 130, 139, *148*, *149*, 188, *201*
Burma, D., 187, *201*, *204*, 280, *288*
Burns, V. W., 159, *201*
Bussard, A., 278, *287*, 385, *401*
Butler, E. A., *471*
Butler, J. A. V., 13, 14, *61*
Butros, J. M., 196, *201*

C

Cabannes, R., 438, *470*
Cairns, J., 126, 127, *148*
Caldwell, P. C., 299, *351*
Callan, H. G., 69, 107, *109*
Cameron, I. L., 196, *202*
Campbell, A., 146, *148*
Campbell, P. N., 355, 362, 363, 373, *400*, *401*
Canellakis, E. S., 13, *61*, 160, 169, 171, 172, 185, *201*, *203*, *204*, 357, 362, *402*
Canellakis, Z. N., 160, 169, *204*
Cantarow, A., *204*
Cantoni, G. L., 357, 358, 360, *401*, *404*
Caputto, R., 57, *61*
Carcassi, V., 449, *470*
Cardini, C. E., 57, *61*
Carey, W., 147, *148*, *149*
Caro, L. G., 346, *351*
Carrier, R. F., 190, *204*
Carrol, E., 3, 4, *63*
Cartledge, J. L., 248, *265*
Case, M. E., 96, *109*
Caspar, D. L. D., 121, *150*, 479, 480, 483, *517*, *518*
Caspersson, T., 272, *287*, 299, *351*
Cassir, R., 176, *205*
Catlin, B. W., 33, *61*
Cavalieri, L. F., 66, 81, 82, 89, 97, 101, 107, *109*, 144, 145, *148*, 183, 184, *201*, *203*
Celano, A., 119, *148*
Ceppellini, R., 418, 444, 445, 446, 448, 449, 463, 464, 469, *470*, *472*
Chamberlin, M., 280, 284, *287*, 389, *401*
Champagne, M., 145, *148*
Chandhuri, N. K., 9, 62
Chantrenne, H., 384, 385, 395, *401*
Chao, F. C., 292, 294, 301, *351*
Chargaff, E., 50, *63*, 230, *269*, 272, *288*
Charlwood, P. A., 416, 418, *470*
Chase, M. C., 72, *109*, 136, *149*, 162, *202*, 210–211, 244, *265*, *266*, 478, *518*
Cheong, L., 219, *265*

Chernoff, A. I., 423, 424, 458, 461, 467, *470*, *474*
Chiozzotto, A., 119, *148*
Chordikion, F., 176, *205*
Chou, S., 156, 174, 175, *202*
Chu, E., 91, *109*, 156, 158, *205*
Chun, E. H. L., 80, *109*
Clark, J. B., 244, *269*
Claude, A., 292, *351*
Clegg, M. D., 408, *470*
Cleland, G. H., 240, *265*
Cleveland, L. R., 67, *109*
Clifton, C. E., 225, *265*
Cobb, M., 232, 233, *268*
Cohen, A., 12, *63*
Cohen, S. S., 8, 9, 10, 12, 46, 55, 56, *60*, *61*, *62*, *64*, 89, *109*, 143, *148*, 156, 158, 162, 170, 174, 194, *200*, 201, 273, 275, *287*, 382, *403*
Cohn, W. E., 272, *287*
Cole, L. J., 166, *203*
Colowick, S. P., 13, *61*
Colter, J. S., 489, *517*
Comly, L. T., 368, 369, *402*
Conley, C. L., 462, 463, *470*
Connel, G. E., 448, *474*
Cook, C. D., 459, *470*
Cook, J. L., 408, *473*
Coppo, A., 119, *148*
Cordero Funes, J. R., *204*
Cordes, E., 356, 358, *400*, *404*
Cordes, S., 119, *148*
Cordonnier, J. K., 416, 442, 443, 454, *473*
Cormick, J., 416, 417, 443, *474*
Cosset, A., 446, *470*
Coughlin, C. A., 218, *265*
Countryman, J. L., 273, *289*
Cousens, S. F., 232, 233, *268*
Cowie, D. B., 178, *204*, 277, *287*, 303, 305, 311, 343, *350*, *351*, 362, 378, *401*
Coy, U., 358, *402*
Craig, D. L., 245, *267*
Craig, L. C., 413, *471*
Crane, H. R., 79, *110*
Cranmore, D., 466, 467, *470*
Crasemann, J. M., 220, 222, *268*
Crathorn, A. R., 13, 14, *61*, 171, 172, 174, 190, *200*, *201*
Creighton, H., 87, *109*
Crick, F. H. C., 26, 38, 46, 49, *64*, 72, 73, *111*, 215, 225, 246, 250, 255, 260, *265*, *269*, 285–286, *287*, 359, 368, 377, 380, 390, *401*, 435, 436, *470*, 510, 511, 512, *517*
Crocker, T. T., 395, 396, 397, *401*
Crosbie, G. W., 6, 7, *61*
Crouse, H. V., 92, *109*
Cullis, A. F., 409, 430, *473*
Cummings, D. J., 142, *149*, 156, 175, *202*
Cunningham, L., 33, *61*

D

Dahl, D., 494, *518*
Dalgliesh, C. E., *401*
D'Amato, F., 225, *265*
Dance, N., 455, *471*
Daniel, V., 368, 369, *401*
Danneberg, P., 9, *62*
Darwin, C., 208, *265*
Das, N. K., 190, 194, *200*, *201*
Davidson, E. A., 17, *61*
Davidson, J. N., 13, *62*, 165, 166, 171, 172, 190, *201*, *202*, *205*, 277, *288*
Davison, P. F., 77, *109*, 122, 123, 125, 126, 130, 139, *148*, *149*, *150*
De, D. N., 92, *109*, 188, *201*
Deasy, C. L., 299, *350*, 354, *401*
Deeley, E. M., 157, 158, 192, *204*
de Garilhe, M. P., 20, 33, *61*
De Giovanni, R., 245, *265*
de Kloet, S., 397, *404*
de la Haba, G. L., 379, *404*
Delamater, E. D., 193, *201*
Delbrück, M., 73, *109*, 208, 241, *267*, *269*, 292, *351*
DeLoze, C., 482, *517*
Dellweg, H., 219, 245, *269*
Demailly, J., Jr., 281, *288*
De Mars, R. I., 225, *265*
Demerec, M., 214, 223, 235, 241, 242, 243, 245, 247, 252, *265*, *266*
De Serres, F. J., 243, 252, *265*, *266*
Destaing, F., 446, *470*
Deutsch, F., 144, 145, *148*
Derreux, S., 385, *401*
De Vries, H., 209, *265*
Dewey, D. L., 194, *202*
De Wolf, S., 9, *62*
Deysson, M., 218, *265*
d'Hérelle, F., 477, *517*

Diamond, L. K., 447, 454, *471*
Dickerson, R. E., 430, *472*
Dickey, F. H., 240, *265*
Dickman, S., 362, *403*
Dieckmann, M., 356, 358, 359, 361, *400, 403*
Dienstbier, Z., 165, 166, *204*
Dinger, R., 347, 348, *351*
Dintzis, H. M., 301, *351,* 374, 375, *401,* 457, *471*
Diringer, R., 382, *402*
Dische, Z., 27, *61*
Dixon, G. H., 448, *474*
Dixon, J., 162, *202*
Djordjevic, B., 79, *109,* 219, 222, *265, 269*
Doctor, B., 361, *401*
Doermann, A. H., 136, 137, *149,* 211, 244, *265*
Dohan, F. C., Jr., 156, 191, *203*
Donalson, J. T., 156, *201*
Donini, P., 119, *148*
Doty, P., 18, 44, *61, 63,* 81, *109, 110,* 144, 147, *150, 151,* 230, 231, *265, 267,* 488, *517, 518*
Doudney, C. O., 173, 176, 178, 179, 180, 187, *201,* 245, *266*
Dounce, A. L., 185, 188, *201, 205*
Drabkin, D. L., 416, *471*
Drauculic, M., 179, 185, *201, 203*
Drasil, V., 196, *202*
Drescher, H., 421, *471*
Drew, R. M., 156, 157, 169, 194, *204*
Dreyer, W. J., 428, *471*
Dreyfus, J. C., 374, 398, 399, *402, 403*
Dubin, D. T., 119, *148*
Duerksen, J. D., 298, 303, 305, 311, *351, 352,* 362, 378, *401*
Dulbecco, R., 244, *269*
Dunn, D. B., 49, *61,* 98, *109,* 120, *149,* 219, *265, 267,* 357, *401,* 488, *518*
Duschinsky, R., 9, *62*
Dutton, R. H., 196, *201*
Dutton, R. W., 196, *201*
Duzer, A., 446, *470*

E

Eady, J. D., 196, *201*
Eason, R., 190, *205*
Ebert, M., 243, *264*
Ebisuzaki, K., 16, *64*
Edelman, I. S., 294, *351*
Edgar, R. S., 96, *109,* 132, 136, *149, 151*
Edington, G. M., 415, 462, 463, *471*
Edmunds, M., 3, *61*
Edwards, J. L., 156, *201*
Efron, M. L., 423, 431, 432, 433, 447, *471*
Eidinoff, M. L., 219, *265*
Eisenstadt, J. M., 187, *201,* 281, *288,* 389, 394, *403*
El-Ani, A. S., 96, *109*
Ellis, F., 156, *203*
Ellis, M. E., 166, *203*
Ellmauer, H., 244, *265*
Elson, D., 298, 303, 305, *351*
Elwyn, D., 7, *61*
Endo, H. W., 245, *265*
Engler, R., 490, *517*
Ephrati-Elizur, E., 220, *269*
Ephrussi-Taylor, H., 125, *149,* 226, *267*
Epstein, H. T., 119, *148*
Erikson, R. O., 159, 163, *201*
Errera, M., 179, *201,* 396, *403*
Erslev, A. J., 450, 451, *470*
Evans, E. A., Jr., 162, 165, 166, *204*

F

Fahmy, M. J., 233, 235, *265*
Fahmy, O. G., 233, 235, *265*
Falcone, G., 193, *201*
Falkow, S., 147, *149, 150*
Fankuchen, I. J., 114, *148*
Fasman, G. W., 482, *517*
Favelukes, G., 363, 369, 370, 371, 372, 379, 386, 397, 398, 399, *400*
Favelukes, S., 370, *403*
Feinendegen, L. E., 169, *201,* 396, *401*
Fessas, P., 442, 445, 447, 453, 454, *471*
Fessenden, J., 367, *401*
Field, E. O., 440, *471*
Fiers, W., 122, *149*
Finamore, F. J., 164, *201,* 279, *287*
Firket, H., 156, *201*
Firshein, W., 196, *201*
Fischer, G. A., 170, *203*
Fischer, R., 243, *265, 267*
Fisher, H., 225, 265
Fisher, J., 380, *403*
Fisher, T. N., 276, *286,* 337, *350*
Fitzgerald, P. A., 454, *471*
Fitz-James, P. C., 155, 158, *206*

Flaks, J. G., 8, 9, 10, 55, *61*, *62*, 162, *201*, 273, *287*
Flynn, F. V., *471*
Ford, C. E., 466, *470*
Foro, F., Jr., 76, *109*, 142, *149*, 161, 183, *201*, 346, *351*
Fosse, R., 227, *265*
Foster, T., 158, 163, *201*
Fottrell, P., 174, *201*
Fox, E., 220, 222, *268*
Fox, M. S., 193, *201*
Fraenkel-Conrat, H., 120, 121, *149*, 240, *266*, 284, 285, *288*, 382, *401*, 478, 479, 483, 484, 485, 486, 487, 488, 489, 490, 491, 492, 495, 497, 498, 499, 500, 501, 504, 509, 512, 513, *517*, *518*, *519*, *520*
Franklin, R. E., 114, *149*, 479, 487, *518*
Franzen, J. S., 170, *201*
Fraser, D., 10, *62*, 117, 120, *149*, *150*, 244, *266*, 273, *287*
Frederiksen, S., 243, *266*
Freese, E., 49, *62*, 96, 97, *109*, 142, *149*, 217, 219, 220, 222, 223, 226, 227, 228, 229, 230, 231, 234, 237, 238, 239, 246, 247, 248, 249, 250, 251, 252, 253, 255, 257, 259, 261, 263, *264*, *265*, *266*, 453, *471*
Freese, E. B., 222, 227, 228, 229, 230, 238, 239, 249, 250, 251, 253, 255, 257, 258, 259, 260, 261, 262, 263, *266*
Freese, H. L., 156, *201*
Freifelder, D., 123, 126, 139, *149*
French, E. F., 312, *351*
Frenster, A., 395, *401*
Fresco, J. R., 247, *266*, 488, *517*
Frick, P. G., 434, *471*
Fricke, H., 242, *266*
Friedewald, W. F., 242, *266*
Friedkin, M., 5, 6, 7, 8, 10, 11, 50, *62*, *64*, 160, *202*
Frisch-Niggemeyer, W., 478, *518*
Frontali, C., 119, *148*
Fuerst, C. J., 119, *151*
Fukai, H., 375, *403*
Fukuhara, H., 281, *288*
Funatsu, G., 486, *519*

G

Gall, J. G., 69, 76, 92, 96, 107, *109*, 159, 187, *202*
Gallant, J., 179, 187, *202*
Gammack, D. B., 423, 424, 438, 441, 446, *471*, *473*
Gamow, G., 285, *287*, 510, *518*
Gard, S., 243, *266*
Garen, A., 428, *471*
Garlick, J. P., 445, *474*
Gaschlbauer, W., 172, *206*
Gay, H., 97, *109*
Gehatia, M., 142, *150*
Gehring-Müller, R., 413, 414, 415, 416, 417, *470*
Geiduschek, E. P., 79, *109*, 185–186, *202*, 227, *266*, 280, *287*
Gelber, B., 160, 188, *206*
Gelfant, S., 191, *202*
Gellert, M., 114, 115, *149*
Genin, C., 281, *288*
Gerald, P. S., 413, 423, 431, 432,, 433, 441, 446, 447, 454, 463, *471*
Giauque, B. G., 157, 169, 194, *204*
Gierer, A., 225, 251, *268*, 382, *401*, 478, 479, 487, 488, 489, 498, 513, *518*
Gilbert, W., 186, *202*, 274, 276, 281, *287*, 315, 326, 329, 344, *351*, 383, 384, 386, *402*
Giles, N. H., 91, 96, *109*, 234, 235, 243, 252, *266*, *267*
Ginoza, W., 487, 490, *518*, *519*
Gish, D. T., 485, 486, *518*, *519*, *520*
Gladstone, L., 280, *289*
Glaess, E., 245, *266*
Glass, B., 241, *266*
Glass, E. A., 218, *266*
Glassman, E., 356, 358, *400*, *401*, *403*
Gnafakis, N., 442, *471*
Goddard, J. W., *204*
Godin, C., 373, *400*
Godman, C. G., 92, *109*, 157, *200*
Goldstein, A., 156, 174, 175, 184, *202*, 375, 377, *401*
Goldstein, D. A., 114, *148*
Goldstein, D. B., 156, 174, 175, *202*
Goldstein, J., 356, *402*, 414, 415, 417, *471*
Goldstein, L., 395, 396, 397, *401*
Goldthwait, D., 311, *352*
Gollub, E. G., 278, *287*
Gordon, J., 358, *401*
Gordon, M. P., 497, *518*
Gosling, R. G., 114, *151*

Gots, J. S., 50, *62,* 278, *287*
Gottschling, H., 219, *266*
Gouttas, A., 453, *471*
Grant, P., 164, *202*
Gratzer, W. B., 406, 408, 416, 418, 428, *470, 471*
Gray, E. D., 13, *62,* 171, 172, *202, 205*
Graziosi, F., 119, *148, 149*
Gregg, J. R., 164, *202*
Green, D., MacD., 239, *266*
Green, M., 10, *62*
Green, M. H., 297, *351*
Greenberg, D. M., 7, 8, *62, 63*
Greenberg, G. R., 9, 16, 50, *62, 64,* 273, *288*
Greengard, O., 362, *401*
Greenough, W. B., III, 459, 467, *474*
Greer, S. B., 218, 222, 231, *266, 269*
Griboff, G., 219, *269*
Griesbach, L., 9, *62*
Griffith, J. S., 285, *287,* 431, *474,* 510, *517*
Gros, F., 174, 176, 179, 186, *202, 203,* 274, 276, 278, 281, *287, 288,* 297, 299, 315, 326, 329, 344, 348, *351, 352,* 363, 382, 383, 384, 385, 386, 394, 399, *401, 402, 403, 404*
Gross, J., 299, *351*
Grossi, L. G., 364, 366, *402, 403*
Grossman, L., 120, *150,* 189, *203*
Gruelich, R. C., 196, *202*
Grumbach, M. M., 88, 89, 91, *110*
Grunberg-Manago, M., 280, *287,* 348, *351*
Guidotti, G., 413, 414, 415, 416, 417, 452, *471, 472*
Gulyas, S., 193, *206*
Guttes, E., 181, 190, *202*
Guttes, S., 159, 181, 190, *202, 204*
Gutz, H., 226, 243, 252, *266*

H

Haagen-Smit, A. J., 299, *350,* 354, *401*
Haas, F. L., 180, 187, *201,* 244, 245, *266, 269*
Hadorn, E., 235, *266*
Hämmerling, J., 395, *402*
Hagerty, G., 24, *63*
Haggmark, A., 3, 4, 5, 9, 50, *60*
Hakala, M. T., 219, *266*
Halberg, F., 166, 191, *200*
Halbrecht, J., 421, *471*
Hall, B. D., 186, *202,* 274, 275, 279, *286, 287, 288,* 310, 315, 347, *351, 352,* 381, 383, 389, *400, 402, 403*
Hall, C. E., 81, 97, *109,* 295, 296, 307, *351*
Hall, D., 186, *205*
Halvorson, H., 362, 378, 384, *402, 404*
Hamaguchi, K., 185–186, *202*
Hamilton, L. D., 125, 126, *149, 150*
Hamilton, M. G., 292, 301, *352*
Hammarsten, E., 3, *62*
Hanawalt, P. C., 89, *110,* 142, *149, 150,* 156, 174, 175, *202, 203*
Handschumacher, R. E., 218, 219, *266*
Hanson, J., 245, *265*
Hanzon, V., 307, *351*
Hardesty, B., *399*
Harding, C. V., 191, *202*
Harm, W., 227, *266*
Harold, F. M., 179, *202*
Harrar, G., 132, *151*
Harris, J. I., 303, *352,* 485, *518*
Harris, J. K., 485, *518*
Hart, R. G., 487, *518*
Hartmann, G., 356, 358, *402*
Hartmann, P. E., 214, *265*
Haschemeyer, R., 479, 488, *518*
Haselkorn, R., 488, *517, 518*
Hasserodt, V., 416, 418, 440, *471*
Haut, W. F., 93, 95, *111,* 195, *206*
Hayashi, M., 275, 276, 279, *287,* 337, *351,* 382, 383, *402*
Hecht, L. I., 13, *62,* 166, 171, *202,* 357, 359, 363, 364, 365, *402*
Hede, R., 123, *149*
Heidelberger, C., 9, *62*
Heiner, R. E., 238, *266*
Helinski, D. R., 215, *266,* 285, *287,* 428, *471*
Hell, A., 396, *403*
Hems, R., 17, *63*
Hendley, D. D., 225, *264*
Hendrickse, R. G., 454, *471*
Henning, U., 285, *287*
Heppel, L. A., 13, 24, 33, *62, 63,* 280, *288*
Herbert, E., 13, *62,* 171, *202,* 357, 362, *402*
Hermodsson, L. M., 307, *351*
Hershey, A. D., 72, *109,* 119, 122, 123, 124, 125, 130, 136, 139, *148, 149, 150,* 162, 188, *201, 202,* 210–211, *266,* 330, *352,* 478, *518*

Hiatt, C. W., 225, *266, 267*
Hiatt, H. H., 14, *61, 62,* 171, 172, 174, 186, *200, 202,* 274, 276, 281, *287,* 315, 326, 329, 344, *351,* 383, 384, 386, *402*
Hidaka, F., 245, *265*
Hieulle, A., 227, *265*
Hill, N. O., 119, *150*
Hill, R. J., 413, 414, 415, 416, 417, *471, 472*
Hill, R. L., 415, 424, 427, *471, 474*
Hillschmann, N., 413, 414, 415, 416, 417, *470*
Hilmoe, R. J., 24, 33, *62, 63*
Hilse, K., 413, 414, 415, 416, 417, *470*
Hinshelwood, C. N., 299, *351*
Hirokawa, R., 362, 363, 364, *402, 403*
Hirota, Y., 225, *266*
Hiruk, C., 245, *265*
Hitzig, W. H., 434, *471*
Hoagland, M. B., 354, 355, 356, 357, 362, 363, 364, 365, 368, 369, 374, *402*
Hobam, G., 413, 414, 415, 416, 417, *470*
Hoch, H., 407, *471*
Hodes, M. E., 230, *269*
Hörlein, H., 422, *471*
Hoff-Jørgensen, E., 164, *202*
Hogness, D. S., 133, *149*
Hokin, L. E., 277, *287*
Hokin, M. R., 277, *287*
Holden, M., 243, *267*
Hollaender, A., 243, *266*
Holley, R. W., 356, 358, 361, *401*
Hollingworth, B. R., 296, 301, 305, 307, *352*
Holloway, B., 354, *402*
Holmes, K. C., 479, 487, *518*
Holzner, J. H., 176, *202*
Hooper, C. W., 125, *150*
Hopkins, J. W., 281, *288*
Hopkins, J., 387, *404*
Hornsey, S., 156, *202*
Horton, B., 444, 448, *471*
Hotchkiss, R. D., 157–158, 193, *201, 202*
Hotta, Y., 159, 163, 164, 171, 174, *202, 205*
Howard, A., 155, 156, 194, *202*
Hoyer, B. H., 294, *350*
Hu, A., 362, 378, *402*
Huang, R. C., 186, *200,* 280, *287,* 397, *401*
Huehns, E. R., 421, 423, 424, 438, 441, 442, 446, 455, *471, 473*
Hülsmann, W. C., 356, 358, 364, 365, 367, *402, 403*
Hughes, W. T., 66, 73, 74, 75, 104, *111*
Hughes, W. L., 138, *151,* 156, 161, 183, *206*
Hughes-Schrader, S., 67, *109*
Huisman, T. H. J., 408, 418, 434, 444, 445, 448, 456, 462, 463, *471, 472, 473, 474*
Hultin, T., 354, 363, 364, 365, 366, 368, 379, *402, 404*
Humphreys, G. K., 7, 8, *62*
Hunt, J. A., 416, 424, 425, 436, 438, 443, 454, *472*
Hunter, A. W. S., 90, 92, *109*
Hurlbert, R. B., 3, *63*
Hurwitz, J., 14, *62,* 280, *287,* 347, 348, *351,* 382, *402*
Huskins, C. L., 90, 92, *109*
Hutchinson, W. D., 440, *474*
Huxley, H. E., 296, 307, *351*

I

Ingraham, L., 125, *149*
Ingram, V. M., 406, 409, 410, 413, 416, 418, 419, 420, 423, 424, 425, 426, 427, 436, 438, 440, 442, 443, 444, 446, 447, 448, 449, 450, 451, 452, 453, 454, 457, 465, *470, 471, 472, 473*
Inman, R. B., 50, *60*
Irvine, R. A., 462, *473*
Ishihama, A., 315, *352*
Ishii, M., 490, *519*
Itano, H. A., 407, 422, 423, 424, 426, 438, 439, 440, 441, 442, 451, 457, 458, *472, 473, 474*
Iwamura, T., 158, *202*
Iwanowsky, D., 477, *518*
Iyer, V. N., 218, 234, *266*

J

Jackson, C. H., 469, *474*
Jacob, F., 145, 146, *149,* 186, 190, *200, 202,* 244, *266,* 272, 274, 277, *287, 288,* 310, 344, *350, 351,* 381, 383, 384, 389, 397, *401, 402, 403,* 453, 464, *472*
Jacob, G. F., 415, 462, 463, *472*
Jaffe, J. J., 13, *61,* 171, 172, *201*
Jalin, T. L., 192, *205*

Jandl, J. H., 467, *469*
Jardetzky, C. D., 166, 191, *200*
Jeener, R., 497, *518*
Jensen, K. A., 234, 240, *266, 267*
Jerne, N. K., 139, *151*
Jesaitis, M. A., 56, *62*
Jim, R. T. S., 424, 427, *474*
Jkeda, Y., 245, *267*
Joklik, W. K., 17, *60*
Jones, A. S., 230, *264*
Jones, O., 366, 390, 391, 392, 393, *403*
Jones, O. W., 282, 283, 284, *288,* 434, 435, 436, *473,* 512, 515, *519*
Jones, R. T., 416, 417, 418, 421, 441, 443, 453, *472, 473, 474*
Jonxis, J. H. P., 374, *402,* 415, 419, 446, 447, 459, *470, 472, 473*
Josse, J., 16, 32, 33, 34, 35, 36, 37, 39, 46, 55, 56, 57, 58, 59, *62,* 162, *202, 203,* 273, *287,* 337, *351*

K

Kaars Sijpesteijn, J. A., 447, *473*
Kaesberg, P., 284, *289*
Kahn, P., 141, *149*
Kaiser, A. D., 32, 33, 34, 35, 36, 37, 39, 46, *62,* 133, 134, 135, *149,* 337, *351*
Kalckar, H. M., 5, 13, *61, 62*
Kalle, G. P., 50, *62*
Kameyama, T., 281, *288,* 348, *351,* 362, 387, 389, 394, *402, 403*
Kanazir, D., 165, 166, *202,* 218, *267*
Kane, M. R., 12, *63*
Kaplan, E., 450, *475*
Kaplan, M. S., 222, *267*
Kaplan, R. W., 244, *265*
Kaplan, W. D., 240, *268*
Karaklis, A., 442, 445, 447, *471*
Karnofsky, D. A., 165, *202*
Karpfel, Z., 196, *202*
Kas, E., 185, *203*
Kaudewitz, F., 225, *267*
Kaufman, E., 225, *267*
Kaufmann, B. P., 97
Kay, E. R., 185, *201*
Kayeyama, T., 187, *201*
Kayser, A., 281, *288*
Keck, K., 10, *62,* 273, *287*
Keighly, G., 299, *350,* 354, *401*
Keir, H. M., 13, *62,* 171, 172, 186, 190, *202, 205*
Kellenberger, E., 98, *109,* 117, 118, 119, 142, *149,* 179, 180, 186, 187, 188, *203*
Kellenberger, G., 86, 87, *109,* 121, 141, *149, 150*
Keller, E., 354, *404*
Keller, E. B., 299, *351*
Kendrew, F. R. S., 430, 431, 432, 433, *474*
Kendrew, J. C., 285, *289,* 430, *472*
Kernot, B. A., 362, *401*
Khorana, H. G., 24, *62*
Kielley, R. K., 13, *62*
Kihara, H., 362, 378, *402*
Kihlman, B., 218, 235, *267*
Kilkson, R., 117, *150*
Kimball, R. F., 157, 187, *203, 204,* 243, *266*
Kingma, S., 424, 434, *473*
Kinosita, R., 156, *205*
Kirby-Smith, J. S., 245, *267*
Kirchner, C. W. J., 252, 264, *267*
Kirk, J., 240, *267*
Kirkland, R. J., 17, *62*
Kirschfeld, S., 219, *269*
Kitani, Y., 96, *109*
Kitazume, Y., 276, *287*
Kjeldgaard, N. O., 192, 195, *203, 205*
Kleinhauer, E., 461, 462, *472*
Kleinschmidt, A., 142, *150*
Klenow, H., 13, 17, *62,* 170, *203, 204,* 243, *266*
Klibansky, C., 421, *471*
Klug, A., 121, *150,* 479, 480, 483, 487, *518*
Kluss, B. C., 190, *204*
Knight, C. A., 485, 486, 488, 489, 493, 494, 497, *517, 518, 519, 520*
Koch, A. L., 156, *201,* 218, *267*
Koch, B., 462, 463, *472*
Koch, G., 119, *150*
Koerner, J. F., 50, 55, 57, *62*
Koler, R. D., 453, *473*
Kolmark, G., 234, 235, 240, *266, 267*
Konigsberg, W., 413, 414, 415, 416, 417, *471, 472*
Konrad, M., 245, *267*
Konzak, C. F., 238, *266*
Kornberg, A., 5, 6, 13, 14, 16, 20, 21, 22, 23, 24, *25,* 26, 27, 29, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45,

46, 47, 48, 49, 50, 52, 55, 56, 57, 58, 59, *60, 61, 62, 63, 64,* 160, 161, 162, *200, 203,* 253, *267,* 273, 280, *287,* 337, *351*

Kornberg, S. R., 16, 46, 55, 56, 57, 58, 59, *62, 64,* 162, *203,* 253, *267,* 273, *287*

Kozinski, A. W., 86, *110,* 140, *150,* 222, *267*

Kramer, G., 482, *518*

Krakow, J. S., 13, *61,* 171, 172, *201*

Kraus, A. P., 462, 463, *472*

Krebs, H. A., 17, *63*

Krevans, J. R., 445, 463, 464, *474*

Krieg, D. R., 239, 244, *266, 267*

Kroger, H., 187, *201, 204,* 280, *288*

Krone, W., 357, *404*

Kruh, J., 374, 398, 399, *402, 403*

Künzer, W., 421, *471*

Kulad, C. G., 186, *202*

Kunitz, M., 19, *63*

Kunkel, H. G., 407, 408, 418, 420, 445, 449, *472*

Kuriki, T., 165, 166, *204*

Kuriki, Y., 164, 165, 166, *203, 204*

Kurland, C. A., 315, 326, 329, *351*

Kurland, C. G., 274, 276, 281, *287,* 305, 344, *351,* 383, *402*

L

Labie, D., 374, *403*

Lacks, S., 363, *402*

LaCour, L. F., 76, *110*

Lagerkvist, V., 360, 361, *400*

Laird, C., 220, 222, *268*

Laite, M. B., 156, *201*

Lajtha, L. G., 156, *203,* 466, *472*

Lamarck, J. B., 208, *267*

Lamborg, M., 355, 386, *402*

Lamfrom, H., 355, 362, 375, 378, 398, 399, *402, 403*

Lamond, W. A., 218, *267*

Lampen, J. O., 5, *63*

Langridge, R., 125, *150*

Lark, C., 195, *203*

Lark, K. G., 156, 158, 165, 166, 167, 169, 170, 173, 178, 179, 180, 181, 182, 184, 186, 187, 190, 191, 192, 195, 196, 197, *203,* 296, *351*

Laskowski, M., 20, 24, *61, 63*

Laskowski, M., Sr., 185, *203*

Latarjet, R., 244, *267*

Lauderbach, A. L., 192, *205*

Lauffer, M. A., 114, 126, *148,* 480, *517*

Laurila, U., 24, *63*

Lawley, P. D., 237, *265*

Lea, D. E., 242, *267*

Leahy, J., 364, 365, 366, 375, *400*

Leavell, B. S., 457, *474*

Legault, R. R., 238, *266*

Lehman, J. R., 253, *267*

Lehman, I. R., 13, 14, 18, 20, 21, 23, 24, 26, 27, 28, 29, 31, 32, 33, 40, 41, 42, 43, 44, 45, 46, 47, 48, 50, 56, 57, *60, 61, 63,* 122, *150,* 160, *200, 203,* 280, *287*

Lehmann, H., 406, 415, 422, 423, 424, 425, 426, 427, 436, 441, 449, 450, 454, 462, 463, *469, 470, 471, 472, 474*

Leloir, L. F., 57, *61*

Lengyel, P., 282, 283, *288,* 390, 391, 392, 394, 400, *400, 402, 404,* 434, 435, 436, *473,* 474, 512, 515, *518, 519*

LePage, G. A., 165, *203*

Lerman, L. S., 119, *150,* 193, *203,* 225, *267*

Lett, J. T., 237, 238, *264*

Leuppold, V., 245, 252, *267*

Levan, A., 218, *267*

Levine, L., 119, 120, *150*

Levinthal, C., 77, 79, *109, 110,* 122, 123, 125, 130, 136, 139, *149, 150,* 161, *203,* 225, *265,* 428, *471*

Lichtenstein, J., 9, *61, 62,* 170, *201,* 273, 275, *287*

Lichtler, E., 13, *62*

Lie, H. G., 455, *473*

Lieberman, I., 5, *62,* 176, *203*

Liebold, B., 415, *470, 473*

Lie-Injo, L. E., 454, 455, *473*

Lima-de-Faria, A., 68, 88, *110*

Lipmann, F., 188, *203,* 356, 357, 358, 363, 364, 365, 366, 367, 368, 370, 371, 373, 379, 380, 394, 399, *402, 403, 404*

Lipschitz, R., 230, *269*

Lissouba, P., 96, *110*

Litman, R. M., 219, 221, 222, 226, *267*

Littauer, U. Z., 368, 369, *401*

Litt, M., 488, *517*

Littlefield, J. W., 80, *109,* 219, *267,* 299, *351,* 488, *518*

Liu, J. C., 423, 424, *470*

Livingstone, F. B., 422, *473*
Lochte, H. L., Jr., 459, 467, *474*
Loeb, M. R., 9, *61*
Løtrup, S., 164, *202*
Loftfield, R., 373, *402*
Logan, R., 277, *288*
Longer, L., 170, *203*
Lorkiewicz, Z., 220, 222, 233, 235, *267, 269*
Lotz, C., 240, *265*
Loutit, J. F., 466, *470*
Loveless, A., 234, 235, 238, *267*
Lowy, P. H., 299, *350*, 354, *401*
Luippold, H. E., 91, 93, *111*, 242, *269*
Luria, S. E., 208, 225, 243, *265*, *267*, 292, *351*
Luzzati, V., 126, *150*, 305, *352*
Lyons, M. E., 88, *110*

M

Maaløe, O., 89, *110*, 142, 143, *149, 150*, 155, 156, 158, 170, 174, 175, 192, 195, *202, 203, 205*, 243, *266*
McCalla, K., 416, 417, 443, *474*
McCarthy, B. J., 179, *200*, 277, *287*, 296, 297, 298, 305, 312, 313, 315, 322, 323, 329, 333, 344, *350, 351*
McCarty, M., 72, *108*
McClary, D. O., 157, *204*
McClintock, B., 82, 87, *109, 110*
McClure, F. T., 277, *288*
McDonald, B. B., 157, 190, *203*
McDonald, M. R., 97, *109*
McElya, A. B., 13, *61*, 173, *200*
McFall, E., 155, *203*, 384, *403*
McGregor, H. C., 69, 107, *109*
MacIver, J. E., 462, 463, *473, 474*
Mackal, R. P., 162, 165, 166, *204*
McKinley, E. B., 243, *265, 267*
McLeish, J., 157, *203*
MacLeod, C. M., 72, *108*
McMaster, R. D., 163, *206*
McNutt, W. S., 3, 4, 5, *63*
McQuillen, K., 297, 299, 305, 312, *352*
Maestre, M. F., 117, *150*
Magrath, D. J., 243, *269*
Maheshwari, N., 186, *200*, 280, *287*, 397, *401*
Mahler, H. R., 10, *62*, 120, *150*, 273, *287*
Main, R. K., 166, *203*
Maley, F., 10, 15, *63*
Maley, G. F., 10, *63*
Mandel, H. G., 497, *519*
Mandel, L. R., 10, *63*
Mandell, J. D., 122, *150*, 330, *352*
Manson, L. A., 5, *63*
Mantsavinos, R., 13, *61*
Marinelli, L. D., 241, *267*
Markham, R., 24, 33, *62, 63*, 493, 497, *519*
Marmur, J., 44, *63*, 81, *110*, 119, 144, 145, 147, *148, 149, 150, 151*, 189, *203*, 230, 231, *265, 267*
Marquardt, H., 241, *267*
Martin, N. J., 409, 410, *473*
Martin, R., 390, 391, 392, 393, *403*
Martin, R. G., 282, 283, *288*, 512, 515, *519*, 434, 435, 436, *473*
Maruyama, Y., 156, 158, 170, 173, 178, 190, 197, *203*
Marvin, D. A., 125, 126, *149, 150*
Mashima, S., 245, *267*
Masson, F., 126, *150*
Mather, K., 90, 92, *110*
Matheson, A. T., 119, *150*
Matsudo, G., 416, 418, 421, *473, 474*
Matthaei, J. H., 282, 283, 284, *288*, 345, 348, *352*, 366, 387, 390, 391, 392, 393, 394, *403*, 434, 435, 436, *473*, 490, 491, 511, 512, 515, *519, 520*
Matthews, R. E. F., 497, *519*
Maximow, A. A., 459, *473*
Mayor, H. D., 119, *150*
Mazia, D., 73, *110*, 197, *200, 203*
Mendel, G., 209, *267*
Mendelsohn, M. L., 156, 161, 191, *203*
Merrill, S., 358, *402*
Meselson, M., 66, 77, 79, 80, 81, 86, 107, *110*, 121, 133, 138, 139, 141, 142, *150, 151*, 161, 183, 186, 189, *200, 203*, 215, *267*, 274, *287*, 310, 344, *350*, 383, *401*
Meyering, C. A., 408, 444, *471*
Meyers, R. W., 480, *519*
Michelson, A. M., 20, *63*
Micou, J., 395, 396, 397, *401*
Midgley, J. E., 277, *287*, 306, 311, 330, 331, 337, 340, *352*
Miescher, F., 272, *288*
Miles, H. T., 215, *267*
Miller, A., 398, *403*

Miller, I. R., 79, *110*
Miller, M. L., 422, *473*
Miller, O. L., 69, 107, *110*
Minagawa, T., 119, 120, *150*
Minckler, S., 157, *204*
Minnich, V., 416, 442, 443, 454, *473*
Mirsky, A. E., 72, *110*, 395, 397, *400*, *401*, *404*
Mitchener, J. W., 462, 463, *473*, *474*
Mitchison, J. M., 157, *206*
Mitra, S., 92, 93, *110*
Mitsui, N., 315, *352*
Miura, K., 245, *265*
Miyake, T., 219, 247, *267*
Modiano, G., 418, 454, *474*
Moldave, K., 364, 366, 367, *401*, *402*, *403*
Monod, J., 190, *202*, 272, 277, 278, *287*, *288*, 344, *351*, 381, 383, 384, 385, 389, 397, *401*, *402*, *403*, 453, 464, *472*
Monroe, R., 366, 370, 380, 394, 399, *403*
Montgomery, T. L., 450, 451, *470*
Montreuil, J., 277, *287*
Montsavinos, R., 171, 172, 185, *201*, *203*
Moore, C. V., 416, 442, 443, 454, *473*
Moore, E. C., *63*
Moore, H. A., Jr., 156, 191, *203*
Mooren, D., 9, *62*
Mordah, J., *204*
Morishima, A., 88, 89, 91, *110*
Morris, A. J., 362, 364, 365, 366, 370, 371, 374, 375, 378, 379, 394, 399, *400*, *403*
Morris, N. R., 170, *203*
Morris, R., 68, *110*
Morrison, M., 408, *473*
Moser, H., 234, 240, *264*
Moses, M. J., 72, 91, 107, *110*
Motulsky, A. G., 453, 454, *473*
Moussear, J., 96, *110*
Mueller, G. C., 194, *205*
Müller, R., 415, *470*
Müller-Eberhardt, V., 418, 445, 449, *472*
Muirhead, H., 409, 430, *473*
Muller, C. J., 419, 420, 421, 424, 434, 446, 447, *470*, *473*
Muller, H. J., 210, *268*
Mundry, K. W., 225, 251, *268*, 498, 499, *518*
Munsch-Peterson, A., 170, *203*
Murakami, W. T., 120, *150*
Murayama, M., 423, 446, 450, *473*
Muzzolini, M., 418, 545, *474*

N

Nakada, D., 175, 180, *203*, 248, *268*
Nakamoto, T., 186, *202*, 280, *287*, *289*, 347, 348, *352*, 382, *404*
Naono, S., 278, *287*, *288*, 385, *401*, *402*, *403*
Narita, K., 483, 485, *517*, *518*
Nasrat, G. E., 240, *268*
Nathans, D., 188, *203*, 364, 365, 366, 367, 368, 370, 373, 380, 394, 397, 399, *403*
Neeb, H., 447, *473*
Neel, J. V., 406, 422, 436, 449, 450, 457, 463–464, 465, *472*, *473*, *474*, *475*
Neidhardt, F. G., 176, 179, *204*
Newmark, P., 480, *519*
Newsome, F. E., 469, *474*
Newton, A. A., 157, 158–159, 193, 195, *204*, *205*
Nicolaieff, A., 126, *150*
Niggli, H., 235, *266*
Nilan, R. A., 238, *266*
Nirenberg, M. W., 282, 283, 284, *288*, 345, 348, *352*, 366, 387, 390, 391, 392, 393, 394, *403*, 434, 435, 436, *473*, 490, 491, 511, 512, 515, *519*, *520*
Nisman, B., 281, *288*, 310, *352*
Niu, C. I., 485, *519*
Niven, J. S. F., 120, *148*
Nixon, H. L., 479, *519*
Nohara, H., 354, *403*
Nomura, M., 121, *150*, 179, *204*, 274, 279, *288*, 298, 310, 315, *352*, 383, *403*
Norcross, F., 296, *351*
North, A. C. T., 114, 125, *150*, 409, 430, *473*
Novelli, G. D., 187, *201*, 281, *288*, 348, *351*, 355, 356, 362, 387, 389, 394, *402*, *403*, *404*
Novick, A., 218, 219, *266*, *268*
Nygaard, O. F., 159, *204*

O

O'Brien, J. A. P., 440, *471*
Ochoa, S., 13, 15, *63*, 187, *201*, *204*, 280, 282, 283, *287*, *288*, 348, *351*, 390, 391, 392, 394, 400, *400*, *402*, *404*, 434, 435, 436, *473*, *474*, 512, 515, *518*, *519*
O'Donnell, J. F., 162, 165, 166, *204*
Oehlkers, F., 210, 218, 245, *268*

Östergren, G., 92, *110*
Ofengand, E., 356, 358, 359, 361, *400*, *403*
Ogata, K., 354, 362, 363, 364, *402*, *403*
Ogur, M., 157, *204*
Okazaki, R., 164, 165, 166, 168, 169, 174, 176, 180, *203*, *204*, *205*
Okazaki, T., 165, 166, 169, 174, 176, 180, *204*, *205*
Okubo, S., 187, *205*, 233, *269*
Okun, L., 220, 222, *268*
Olive, L. S., 96, *109*, *110*
Oliver, R., 156, *203*, 466, *472*
Olson, M., 362, 374, *403*
Omori, S., 362, 363, 364, *402*, *403*
Oota, Y., 315, *352*
Opara-Kubinska, Z., 220, *269*
Ord, M. G., 165, 166, *204*
Orgel, A., 225, 255, 260, *265*, *268*, 285, *287*, 510, *517*
Ortiz, P. J., 348, *351*
Osawa, S., 315, *352*, 357, *403*
Osgood, E. E., 453, *473*
Otaka, E., 315, *352*
Ove, P., 176, *203*
Overgaard-Hansen, K., 170, *204*
Overman, R. R., 469, *474*

P

Padieu, P., 374, *402*
Paigen, K., 121, 133, *151*, 298, *352*
Painter, R. B., 156, 157, 169, 194, *201*, *204*, 396, *401*
Palade, G. E., 292, *352*, 362, 378, *404*
Paladini, A. C., 57, *61*
Papaspyrou, A., 442, 454, *471*
Pappas, G. D., 97, 104, 105, *110*
Pardee, A. B., 155, 159, 160, 174, *200*, *204*, 219, 221, 222, *267*, 277, 292, 298, 343, *352*, 383, 384, *403*
Parizek, J., 165, 166, *204*
Parkins, G. M., 237, 238, *264*
Partridge, C. W. H., 243, *266*
Paul, J., 13, 17, *64*, 172, 173–174, 176, 177, *201*, *204*, *206*
Pauling, L., 407, 409, 422, 457, *473*
Payne, R. A., 444, 448, *471*
Pearce, J. D., 196, *201*
Pelc, S. R., 76, *110*, 155, 156, *202*
Perry, R. P., 396, *403*
Perutz, M. F., 409, 430, *473*
Petermann, M. L., 292, 301, *352*
Peters, T., 362, *403*
Peterson, E. A., 408, *473*, 489, *519*
Phear, E. A., 7, *63*
Phillips, D. C., 430, *472*
Pinkerton, T. C., 130, 139, *151*
Pirie, N. W., 478, *517*
Pizer, L., 382, *403*
Plaut, W. S., 73, *110*, 396, *401*
Plever, E., 9, *62*
Plough, H. H., 231, *268*
Pogo, A. O., *204*
Pomper, S., 245, *268*
Popper, H., 176, *202*
Portier, A., 438, 446, *470*
Poschkis, K. E., *204*
Potter, R. L., 12, *63*, 165, 166, *204*
Potter, V. R., 12–13, 51, *61*, *62*, *63*, 166, 171, 172, 173, *200*, *202*
Poulik, M. D., 407, *473*
Pouyet, J., 18–19, *63*
Prankerd, T. A. J., 451, *470*
Pratt, D., 222, 253, *267*, *268*
Pratt, E. A., 56, 57, *63*
Preiss, J., 356, 357, 359, *403*
Prescott, D., 396, *403*
Prescott, D. M., 156, 187, 190, *204*
Prestidge, L. S., 174, *204*, 298, 343, *352*
Price, C. C., 232, *268*
Prins, H. K., 408, *473*
Prusoff, W. H., 170, *204*
Punt, K., 444, 445, *472*
Putnam, F. W., 139, *150*, 240, *268*

Q

Quastler, H., 191, 193, 194, *204*, *205*

R

Raacke, I. D., 355, 357, 358, 362, *403*
Rabinowitz, M., 362, 374, 380, *403*
Racker, E., 5, *63*
Radding, C. M., 40, 41, 42, 43, 44, 45, *63*
Ramachandran, L. K., 479, 485, *517*, *518*, *519*
Ranney, H. M., 426, 436, 454, *470*, *473*
Ransome, S., 454, *471*
Raper, A. B., 415, 438, 441, 462, 463, *472*, *473*
Rapoport, J. A., 210, 234, 235, 240, *268*
Razell, W. E., 24, *62*

Read, J., 242, *269*
Reddi, K. K., 347, 348, *352*
Reichard, P., 3, 4, 5, 9, 50, *60, 62, 63,* 160, 169, *204*
Reiner, B., 233, 237, 238, *268*
Remy, C. N., 5, *63*
Remy, W. T., 5, *63*
Rhinesmith, H. S., 409, 410, *473*
Rho, J. H., 395, 396, *403*
Rich, A., 114, 125, *150,* 449, 461, 462, *473*
Rich, M. A., 219, *265*
Richards, B. M., 157, 158, 192, *204*
Richards, J., 13, *62,* 171, 172, *202, 205*
Rigas, D. A., 453, *473*
Riley, M., 277, *288,* 383, *403*
Ripley, S., 354, *402*
Ris, H., 72, 96, 97, *110*
Rissebrough, R. W., 186, *202,* 274, 276, 281, *287,* 315, 326, 329, 344, *351,* 383, *402*
Ritter, D. B., 294, *350*
Rixon, R. H., 192, *206*
Rizet, G., 96, *110*
Roberts, D., 5, *62*
Roberts, R. B., 178, 179, 186, *200, 204,* 277, *287, 288,* 293, 297, 298, 299, 301, 303, 305, 309–310, 311, 312, 313, 315, 333, 344, *350, 351, 352,* 362, 363, 378, *401, 403*
Robinson, A. R., 422, 450, *473, 475*
Robinson, E. A., 423, 424, 426, 438, 439, 441, 442, *472*
Robson, J. M., 210, 234, 235, *264*
Rogers, S., 219, *268*
Rolfe, R., 143, *150*
Roll, P. M., 3, 4, *63*
Roller, A., 140, *150*
Rosa, J., 398, 399, *402*
Rose, I. A., 3, *63*
Rosenberg, B. H., 66, 81, 82, 89, 97, 101, 107, *109,* 144, 145, *148,* 183, 184, *201, 203*
Rosenkranz, H. S., 107, *108*
Rosenthal, S. M., 119, *148*
Ross, W. C. J., 232, *268*
Rosseels, J., 497, *518*
Rossman, M. G., 409, 430, *473*
Rotherham, J., 165, 166, *204, 205*
Rothman, F., *471*
Rownd, R., 147, *150*
Rubenstein, I., 116, 117, 120, 123, 124, *150*
Rucknagel, D. L., 406, 422, 436, 449, 450, 461, 467, *474*
Rudloff, V., 413, 414, 415, 416, 417, *470*
Rudner, R., 189, *205,* 219, 223, 252, 263, *268*
Rueckert, R. R., 194, *205*
Rusch, H. P., 159, *204*
Rushizky, G. W., 489, 497, *519*
Russell, E., 363, 371, 372, 379, 386, 398, *400*
Rutberg, L., 3, *63*
Ryan, F. J., 175, *203,* 210, 248, *268*
Ryter, A., 117, 118, 119, *149*

S

Sable, H. Z., 12, *63*
Sadron, C., 145, *148*
Saeki, R., 375, *403*
Saluste, E., 3, *62*
Salzberger, M., 461, *470*
Sams, J., 241, 243, *265*
Sarkar, N. K., 185, *201, 205*
Sarvella, P., 68, *110*
Saska, J., 196, *202*
Sato, C., 395, *404*
Sato, G. H., 139, *151*
Sato, J., 375, *403, 404*
Sauerbier, W., 240, *268*
Sauerland, H., 92, *110*
Sax, K., 91, 92, 93, *110,* 231, 241, 242, *268*
Scarano, E., 10, *63*
Schaad, J. D. G., 444, 445, *472*
Schachman, H. K., 29, 40, 41, 42, 43, 44, 45, *63,* 292, *351, 352,* 487, *517*
Schachter, E., 126, *148*
Schachtschabel, D., 357, *404*
Schaechter, M., 142, *149,* 155, 156, 175, 192, 195, *202, 203, 205*
Schaeffer, P., 146, *149*
Schaffer, F. L., 478, *519*
Schairer, M. V., 76, *111*
Schapira, G., 374, 398, 399, *402, 403*
Scheimer, J., 9, *62*
Schell, P., 228, *265*
Scherbaum, O. H., 192, *205, 206*
Schildkraut, C. L., 81, *110,* 144, 147, *150, 151,* 230, *267*

Schlessinger, D., 274, 281, *288*, 296, 297, 299, 301, 307, 347, *352*, 383, 386, 394, 399, *404*
Schlesinger, S., 12, *63*, 165, 166, *204*
Schneider, J. H., 176, *205*, 395, 396, *403*
Schneider, M. J., 248, *268*
Schneider, W. C., 165, 166, *204*, *205*
Schnek, A. G., 408, 420, *474*
Schramm, G., 225, *268*, 382, *401*, 478, 479, 480, 482, 485, 486, 487, 489, 490, 498, *517*, *518*, *519*
Schroeder, W. A., 408, 409, 410, 416, 417, 418, 420, 421, 440, 441, 443, 453, *469*, *470*, *472*, *473*, *474*
Schuitzer, R. J., 9, *62*
Schulman, H. M., 279, *288*
Schumacher, G., 480, *519*
Schuster, H., 225, 227, 228, *268*, *269*, 479, 497, 498, 499, 513, *519*, *520*
Schwartz, D., 82, 96, *110*
Schwartz, H. C., 424, 427, *471*
Schwartz, I. R., 437, 438, 439, 441, 450, 451, *470*
Schweet, R. S., 355, 356, 358, 361, 362, 363, 364, 365, 366, 367, 369, 370, 371, 372, 374, 375, 378, 379, 386, 394, 397, 398, 399, *400*, *401*, *403*, *404*
Schweigert, B. S., 3, *63*
Schwerdt, C. E., 478, *519*
Scopes, A. W., 159, *206*
Scott, D. B. M., 156, 158, *205*
Scott, E. M., 431, *474*
Scott, J. E., 191, *200*
Scott, J. F., 363, 364, 365, *402*
Scott, S. S., 166, *205*
Sechaud, J., 117, 118, 119, *149*
Seed, J., 193, *205*
Seeds, W. E., 114, 125, *150*, *151*
Setlow, R. B., 142, *151*, 184, *205*
Setlow, S. K., 184, *205*
Shapiro, H. S., 50, *63*, 230, *269*
Shelton, J. B., 416, 417, 443, *474*
Shelton, J. R., 416, 417, 443, *474*
Sherman, F. G., 191, 194, *204*, *205*
Shimura, K., 362, 375, *403*, *404*
Shooter, E. M., 423, 424, 438, 441, 442, 445, 446, 455, *471*, *473*, *474*
Shooter, K. V., 13, 14, *61*, 171, 172, 174, 190, *200*, *201*, 455, *471*
Shore, V. C., 430, *472*
Shugar, D., 20, *63*, 245, *268*
Sibatani, A., 397, *404*
Siegel, A., 490, 500, *519*
Siekevitz, P., 12–13, *63*, 354, 362, 378, *404*
Silva, D. M., 494, *518*
Silvestroni, E., 418, 447, 449, 454, *474*
Simkin, J., 378, *404*
Simmons, N. S., 125, *149*
Simms, E. S., 5, 13, 14, 20, 21, 23, 24, 26, 27, 29, 31, 32, 33, 46, 47, 48, 50, *60*, *61*, *62*, *63*, 160, *200*, *203*, 280, *287*
Simon, E. H., 79, *110*, 161, *205*
Singer, B., 120, *149*, 479, 487, 488, 489, 490, 495, 497, *517*, *518*, *519*
Singer, K., 461, *474*
Singer, L., 461, *474*
Singer, M., 357, *404*
Singer, M. F., 24, *63*
Singer, R., 461, *470*
Singer, S. J., 407, 422, 439, 440, 442, 457, *472*, *473*, *474*
Siniscalco, M., 449, *470*
Sinsheimer, R. L., 20, 26, 27, 38, 54, *64*, 119, 121, 122, 139, *149*, *151*, 253, *268*
Sisken, J. E., 156, 187, *205*
Skinner, E. R., 445, *474*
Skoda, J., 165, 166, *204*
Slayter, H. S., 295, 296, 307, *351*
Sluyser, M., 368, *400*, *401*
Smellie, R. M. S., 13, 17, *62*, *64*, 171, 172, 173–174, 186, 190, *202*, *205*, *206*
Smith, C. L., 157, 193, 195, *205*
Smith, D. B., 409, 413, *475*
Smith, E., 392, *404*
Smith, E. L., 435, 436, *474*
Smith, E. W., 438, 439, *474*
Smith, G. M., 454, *470*
Smith, H. W., 243, *269*
Smith, J. D., 49, *61*, *63*, 98, *109*, 120, *149*, 219, 222, *265*, 357, *401*, 493, 497, *519*
Smith, K., 358, *404*
Smith, M. S., 50, 55, 57, *62*
Smith, R., 170, *205*
Smithies, O., 407, 448, *474*
Sneoke, N., 231, *265*
Sobels, F. H., 240, *268*
Sober, H. A., 408, *473*, 489, *519*

Somerville, R. L., 9, 16, 50, *62*, *64*, 273, *288*
Spahr, F., 384, 386, *402*
Spahr, P. F., 305, 306, *352*, 357, *401*
Spencer, M., 126, *150*, 215, *268*
Speyer, J. F., 282, 283, *288*, 390, 391, 392, 394, 400, *400*, *402*, *404*, 434, 435, 436, *473*, *474*, 512, 515, *518*, *519*
Spiegelman, S., 179, 186, 187, *200*, *202*, *205*, 274, 275, 276, 279, 280, 281, *287*, *288*, 303, 305, 310, 311, 315, 337, 347, *351*, *352*, 362, 378, 382, 383, 384, *401*, *402*, *403*, *404*
Spilman, W., 147, *148*, *149*
Spinelli, V., 247, *269*
Sprinson, D. B., 7, *61*
Srinivason, B. D., 191, *202*
Stacey, K. A., 232, 233, *264*, *268*
Stadler, L. J., 210, 245, *268*, *269*
Staehelin, M., 240, *268*, 497, 498, *518*, *519*
Stahl, F. W., 66, 77, 79, 80, 81, 96, 107, *110*, 133, 138, 142, *150*, *151*, 161, 183, 189, *203*, 215, 220, 222, 243, 244, *265*, *267*, *268*
Stamatoyannopoulos, G., 445, 447, *471*
Stanier, R. Y., 292, *352*
Stanley, W. M., 478, 485, 486, 493, 497, *518*, *519*, *520*
Stanners, C. P., 155, 156, 193, *205*, *206*
Steffensen, D., 96, *110*
Steinberg, D., 354, 373, *400*, *404*
Steiner, R. F., 225, *264*, *269*
Stent, G. S., 73, *109*, 119, 139, *151*, 155, 180, *203*, *205*, 222, 244, *268*, *269*, 384, *403*
Stephenson, M. L., 354, 355, 357, 358, 359, 363, 364, 365, *402*, *404*
Stern, C., 87, *111*
Stern, H., 158, 159, 163, 164, 171, 174, *201*, *202*, *205*
Stevens, A., 280, *288*, 382, *404*
Stevens, M. A., 243, *269*
Stocken, L. A., 165, 166, *204*
Stohlman, F., Jr., 466, *474*
Stone, L. H. A., 90, *110*
Stone, W. S., 243, 244, *269*
Storck, R., 186, *205*, 275, *288*, 310, 347, *352*
Strandberg, B. E., 430, *472*
Strauss, B., 187, *205*
Strauss, B. S., 233, 234, 239, 240, *269*
Streisinger, G., 131, 132, *151*
Strelzoff, E., 96, *111*, 221, 223, 252, 263, *269*
Stretton, A. O. W., 410, 418, 419, 444, 447, 449, 450, 451, 452, 453, 457, *472*
Strominger, J. L., 166, *205*
Studier, F. W., *151*
Stulberg, M. P., 355, 356, *404*
Sturgeon, P., 422, *472*
Sueoka, N., 44, *64*, 80, *111*, 161, 183, 189, *205*
Sugino, N., 165, *205*
Sugino, Y., 165, *205*
Sugiyama, T., 488–489, 497, *519*
Suit, H. D., 467, *474*
Sullivan, E., 416, 454, *470*
Sunderland, N., *157*, *203*
Suskind, S. A., 179, 187, *202*
Suzuka, I., 362, *404*
Swanson, C. P., 245, *269*
Swartz, M. N., 38, 39, 45, 47, 49, 50, *64*
Swenson, R. T., 424, 427, *471*, *474*
Swift, H. H., 157, *205*
Swift, J., 160, 188, *206*
Szybalski, W., 79, *109*, 193, *201*, 218, 219, 220, 222, 233, 234, 235, 257, *265*, *266*, *267*, *269*

T

Takahashi, T., 363, 364, *402*
Takahashi, W. N., 490, *519*
Takanami, M., 364, 366, 367, *404*
Takats, S., 164, *205*
Takemura, S., 227, *269*
Tal, M., 303, *351*
Tamm, C., 230, *269*
Taylor, A. L., 146, *151*
Taylor, J. H., 66, 72, 73, 74, 75, 76, 80, 82, 83, 85, 88, 89, 91, 92, 93, 95, 96, 97, 98, 99, 104, 107, *110*, *111*, 138, *151*, 156, 157, 161, 163, 169, 183, 195, *206*
Taylor, L. S., 241, *267*
Tener, G. M., 24, *62*
Tessman, E. S., 119, *151*, 244, *269*
Tessman, I., 119, *151*, 227, *269*
Terasima, T., 159, 192, 194, *206*
Thoday, J. M., 91, *111*, 242, *269*
Thomas, C. A., Jr., 119, 122, 123, 124, 128, 129, 130, 138, 139, *148*, *150*, *151*

Thomas, E. D., 459, 467, *474*
Thompson, L., 12, *63*
Thompson, L., 165, 166, *204*
Thompson, R. B., 462, 463, *473*, *474*
Thorup, O. A., 457, *474*
Thrasher, J. D., 196, *202*
Tijima, T., 225, *266*
Till, J. E., 155, 156, 193, *205*, *206*
Timasheff, S. N., 305, *352*
Timofeeff-Ressovsky, M. W., 241, *269*
Tomlin, P. A., 222, *267*
Tissières, A., 274, 281, *288*, 296, 297, 299, 301, 306, 307, 348, *352*, 383, 386, 387, 394, 399, *404*
Tobita, T., 373, *404*
Tocantins, L. M., 438, 439, 441, 450, 451, *470*
Tolarico, M., 10, *63*
Tolmach, L. J., 159, 192, 193, 194, *203*, *206*
Toschi, G., 307, *351*
Trager, L., 219, *269*
Trautner, T. A., 38, 39, 45, 47, 49, 50, *64*, 136, *151*
Treffers, H. P., 247, *269*
Trobert, J. V., 438, 439, *474*
Tsevremis, H., 453, *471*
T'so, P., *151*, 292, 294, 301, *351*, *352*, 395, *404*
Tsujita, A., 284, 285, *288*, 484, 485, 486, 487, 491, 494, 499, 500, 501, 504, 509, 512, 513, *518*, *519*, *520*
Tung, J., 93, 95, *111*, 195, *206*
Turnbull, E. P. N., 415, 459, *474*
Turner, J. F., 17, *62*
Tuttle, A. H., 469, *474*
Twort, F. W., 477, *520*

U

Uhlig, H., 485, 486, *517*
Utsumi, K. R., 157, *206*

V

Valentine, W. N., 449, *474*
Vallisneri, E., 418, 454, *474*
Van Bibber, M. J., 16, *60*, *61*
Van Tubergen, R. P., 142, *151*
Van Vunakis, H., 119, 120, *150*
Vaughan, M., 373, *404*
Veldee, S., 490, 492, 509, *518*, *520*
Vella, F., 442, 454, *474*
Vendrely, C., 72, *109*, 157, *206*
Vendrely, R., 72, *109*, 157, *206*
Verly, W. G., 156, *201*
Vielmetter, W., 225, 227, 251, *269*, 498, 499, 513, *520*
Vincent, W. S., 276, *287*, *289*, 336, 337, *352*
Vinograd, J., 77, *110*, 292, 294, 301, *351*, *352*, 374, *401*, 416, 418, 440, 441, 442, 453, *471*, *472*, *474*
Vischer, E., 272, *288*
Visser, H. K. A., 459, *472*
Vogt, M., 218, *269*
Volkin, E., 56, *64*, 164, 179, *200*, *201*, 273, 275, 276, 278, 279, *286*, *287*, *288*, *299*, 336, 337, *352*, 386, *400*
von der Decken, A., 364, 365, 366, 368, *404*
von Ehrenstein, G., 363, 364, 365, 366, 370, 371, 379, 380, 394, 399, *403*, *404*

W

Wacker, A., 219, 245, *269*
Wahba, A. J., 7, *64*, 390, *400*
Wake, R. G., 50, *60*
Wakonig, T., 92, *110*
Wales, M., 459, 467, *474*
Walker, J., 415, 459, *474*
Walker, P. M. B., 157, 158, 192, *204*, *206*
Wallenius, G., 407, 408, 418, 420, 449, *472*
Waller, J.-P., 303, *352*
Warner, R., 187, *201*
Warner, R. C., 19, *64*
Warren, W., 311, *352*
Watson, H. C., 276, 285, *289*, 430, 431, 432, 433, *472*, *474*
Watson, J. D., 26, 38, 46, 49, *64*, 72, 73, *111*, 179, 186, *202*, *204*, 215, 246, *269*, 274, 281, *287*, 296, 298, 301, 307, 315, 326, 329, 344, *351*, *352*, 383, 384, 386, *402*
Watts-Tobin, R. J., 250, 255, *265*, 285–286, *287*, 377, 390, *401*, 435, 436, *470*, 511, 512, *517*
Weatherall, D. J., 441, 442, 443, 446, 454, *474*
Weber, E., 485, 486, *517*
Weber, G., 422, *471*

Webster, G. C., 364, 365, 366, 370, *404*
Weed, L. L., 10, *61*
Weigle, J., 80, 86, 87, *109, 110*, 121, 133, 141, *149, 150, 151*, 244, *269*
Weill, J., 187, *201, 204*
Weill, J. D., 280, *288*
Weinblum, D., 245, *269*
Weinfeld, H., 3, 4, *63*
Weisblum, B., 361, *400*
Weiss, S. B., 186, *202*, 280, *287, 289*, 347, 348, *352*, 382, 397, *404*
Weissman, J. M., 13, 17, *64*
Weissman, S. M., *62*, 171, 172, 173–174, *202, 206*
Welch, A. D., 218, 219, *266*
Weliky, N., 418, 421, *473*
Wells, I. C., 407, 422, 457, *473*
Went, L. N., 462, 463, *473, 474*
Wertheimer, S. A., 76, *109*, 142, *149*, 161, 183, *201*
Westergaard, M., 234, 235, 240, 252, *266, 267, 269*
Weygand, F., 219, *269*
Wheby, M., 457, *474*
Wheeler, J. T., 445, 463, 464, *474*
White, F. H., Jr., 482, *520*
White, J. C., 416, 459, 461, *474*
Whitfeld, P., 489, *520*
Whitfield, J. F., 192, *206*
Whitmore, G. F., 193, *206*
Wieder, C. M., 225, 251, *269*
Wilber, P. B., 12, *63*
Wildman, S. G., 490, *519, 520*
Wildy, P., 157, 158–159, 193, 195, *204, 205*
Wilkins, M. H. F., 114, 125, 126, *149, 150, 151*, 215, *269*, 308, *352*
Will, G., 409, 430, *473*
Williams, R. C., 117, 121, *149*, 487, 489, *518*
Williams, W. J., 416, 442, 443, 454, *473*
Williamson, D. H., 159, *206*
Williamson, M. B., 172, *206*
Wilson, H. R., 125, *149, 150*
Wilson, S., 409, 413, *475*
Wimber, D. R., 191, *205*
Wintrobe, M. M., 466, *475*
Witkin, E. M., 187, *206*, 225, 244, 245, *269*
Wittman-Liebold, B., 413, 414, 415, 416, 417, *470*
Wittmann, H. G., 480, 482, 484, 486, 493, 494, 501, 502, 507, 512, *518, 520*
Witz, J., 305, *352*
Woese, C. R., 511, *520*
Wolf, J., 418, 445, 449, *472*
Wolf, M. K., 225, *265*
Wolff, S., 91, 93, *111*, 242, *269*
Wollman, E. L., 145, 146, *149*
Woodard, J., 160, 188, *206*
Wood, W., 388, 394, *404*
Wood, W. B., 283, 284, *289*
Woods, P. S., 66, 73, 74, 75, 76, 104, *111*, 138, *151*, 156, 161, 183, *206*
Woods, R. D., 479, *519*
Work, T., 373, *400*
Wyatt, G. R., 10, 46, 56, *64*
Wyss, O., 244, *269*

X

Xefteri, E., 453, *471*

Y

Yamamoto, N., 225, *269*
Yamazaki, H., 284, *289*
Yanofsky, C., 215, *266*, 285, *287*, 428, *471*
Yarmolinsky, M. B., 379, *404*
Yates, H. B., 157, *206*
Yčas, M., 276, *287, 289*, 336, 337, *352*, 511, *520*
Yehuda, M., 461, *470*
Yoshida, A., 373, *404*
Youcis, P., 156, *201*
Young, I. E., 155, 158, *206*
Young, J., 485, 486, *520*

Z

Zachau, H., 357, *404*
Zahn, R. K., 142, *150*
Zalokar, M., 396, *404*
Zamecnik, P. C., 272, *289*, 299, *351*, 354, 355, 357, 358, 359, 363, 364, 365, 369, 386, *402, 404*
Zamenhof, S., 219, 220, 231, 233, 237, 238, *266, 268, 269*
Zech, H., 490, *520*
Zetterberg, G., 235, *269*
Zeuthen, E., 164, 192, *202, 206*
Zichichi, M. L., 86, 87, *109*, 121, 141, *149, 150*
Zillig, W., 357, *404*, 480, 482, *519*

Zimmer, K. G., 241, *269*
Zimmerman, S. B., 13, 14, 16, 46, 48, 50, 55, 56, 57, 58, 59, *61, 62, 63, 64,* 162, *203,* 253, *267,* 273, *287*
Ziporin, Z. Z., 179, *202*
Zobel, R., 126, *148*
Zogerman, J., *204*
Zubay, G., 296, 307, 308, *351, 352*
Zuelzer, W. W., 422, 450, *473, 475*

# SUBJECT INDEX

A

Acridine dyes as mutagens, 224–225
Adenosine triphosphate-generating system, 14
Adenylate kinase, 12–13
Alkylating agents, 231–237
Amino acyl-RNA, 355–356
2-Aminopurine, 222–225
Analogs, base, 46–50
  substitution in DNA synthesis, 46–50, 219–227
Autoradiography, 73–77, 122–128
  molecular, 123–124, 127

B

Bacteria,
  cytology of, 142–147
Bacteriophage, 113–141
  distribution of parental DNA to progeny, 138–140
  $\phi$X174, 119–122
  ghosts, 115
  HETS, 136
  maturation, 117–119
  organization of DNA, 114–117
  recombination in T2 and T4, 132–133
  structure of heterozygotes, 136–138
  transfer of parental DNA, 139–141
  T-2, 114–119, 123–125
  turnover of RNA in cell, 278–279
Base analogs (see analogs)
Bromodeoxyuridine,
  substitution for thymidine in healing chromosomal lesions, 94
5-Bromouracil, 48–50, 220–222
  incorporation in copolymer, 49

C

Carbon-14 as label for DNA, 18, 73
Chloramphenicol, effects on ribosome synthesis, 333
Chloromercuribenzoate, 55
Chromatid, 67–68
  sister exchangers, 82–85
Chromatography, 315–316
  separation of hemoglobins, 406–407
Chromonema, 68
Chromosome replication and protein synthesis, 89–90
Chromosomes, 65–108
  axial element, 72
  duplication, 73–77
  exchanges, 82–86
  giant, 68–72
  lampbrush, 69–72, 107
  labeling with tritium, 73–77
  mechanism of healing, 93–96
  models of, 96–108
  replication of, 72–82
  reproduction of, 89–92
  reunion of breaks, 92–96
  sequences in duplication, 88–89
  structure, 67–72
    of bacterial, 143–144
  types of changes produced by radiation, 90–92
Cistron, 214
Coding, genetic, 282–286, 390–395
    evidence for triplet, 284–285
      from mutant hemoglobins, 434–435
      from TMV, 509–517
Colchicine and chromatid exchanges, 83
Complementation, 214
Copolymers, 42

D

Deamination by nitrous acid, 225–227
Deoxyriboaldolase, 5
Deoxycytidine, 3
Deoxycytidine triphosphatase, 50
Deoxycytidylate hydroxymethylase, 11–12
Deoxyribonucleic acid (DNA),
  base alterations, 216
  breakage by shear forces, 122, 130–131
  change in state during replication, 184–185
  changes in amount over the cell cycle, 154–158
  control mechanisms and nucleotide pools, 159–160, 162–165
  d-AT polymer, 40–45

d-GC polymer, 45–46
density-gradient centrifugation of, 77–79
"melting temperature," 231
molecular weights of bacterial DNA, 145
nearest neighbor analysis, 32–40
number of chains in molecule, 81, 125–128, 144
organization in phages T2 and T4, 114–117, 120
organization into chromosomes, 96–108, 113–148
phage $\phi$X174, 121–122
precursors of, *see* Deoxyribonucleotides
primer, 25–27, 161, 180
polymerase, *see* Polymerase
replication, 77–82
  semiconservative, 74–80, 161
  sequence of, 188–190
single chain regions or weak points, 119, 128
synthesis, 17–60
  base composition of product, 29–32
  control mechanisms, 160–200
  dissociation from cell division, 197
  effects of injury, 191
    of nutrient environment, 194–196
    of protein synthesis, 173–180
    of radiation, 192–194
    of temperature, 192–194
  limited reaction, 23–24
nature of primer, 25–27
    of product, 27–29
  nearest neighbor analysis of product, 32–40
  physical state of DNA in controlling, 180–183
  relation to the cell cycle, 155, 175–180, 190–192, 197
  replacement of natural bases by analogs, 48
  requirements for cell free system, 20
  simultaneous RNA synthesis, 186–187
  site in the cell, 190
  structural stabilization by protein, 188
  variations in the cycle, 191–197
transfer and recombination in lambda phage, 141
transfer experiments with BUDR, 140
  with *E. coli,* 142
Deoxyribonuclease, 28
  effects on chromosomes, 69
Deoxyribonucleoside,
  monophosphate kinase, 12, 16
Deoxyribonucleotide pools, in anthers of *Lilium* and *Trillium* anthers, 163–164
  in bacteria, 166–169
Deoxyribonucleotides, 1–17
  phosphorylation of, 12–17
  synthesis of, 3–12, 163–172
Deoxyuridine triphosphatase, 50
Deoxyuridylate, synthesis of thymidylate from, 6–8
Depolymerization of DNA, 238
Depurination of DNA, 230, 237–240
Diesterase, splenic, 34
Dyes, mutagenic effects of, 224, 225

E

Electrophoresis, 408
Enzyme induction, 277, 343–344
Enzymes,
  amino acid activating, 354–356
  phosphorylating in control of DNA synthesis, 171
  ribosomal, 303–305
Episomes, bacterial, 147
Exchanges,
  chromatid, induced by radiation, 90–92
  half-chromatid, 92
  single and twin in chromosomes, 83
  sister-chromatid (*see* Chromatid, exchanges)

F

Fingerprints of hemoglobin, 410–412
5-Fluorodeoxyuridylate, 9, 93–94
Folic acid, 6–12
Formaldehyde as a mutagen, 240

G

$\beta$-Galactosidase,
  gene for, 281
  induction of, 277, 281

Gene for beta-galactosidase, 281
Gene mutations, 212–214
Glucosyl transferases, 57–58
Glucosylation of 5-hydroxymethylcytosine, 56–59
Gradient centrifugation,
  cell components, 273–274, 316–321
  DNA, 77–80

H

Half-chromatids, 67, 99
Hemoglobin-$A_2$, 418–419
Hemoglobin-$A_3$, 420–421
Hemoglobin-F, 415, 418, 421
Hemoglobins, 405–469
  abnormal, 421–434
  amino acid substitution in variants, 428–434
  control mechanism in synthesis, 456–461
  dissociation and recombination, 440–442
  genetics of, 436–438, 444–447, 462–465
  mode of assembly of molecule, 439–440
  sequence of amino acids in the chains, 419–420
  structure of, 409–415
Homopolymers, 54
  of DNA, 54
Hydrazine, reaction with DNA, 227
Hydrogen bonds in DNA, 81
Hydrogen ion concentration (pH), mutagenic effect, 230
Hydroxylamine, mutagenic effect of, 227–230
  reaction with bases of DNA, 227–228
    with bromouracil, 229
    with cytosine, 228–229
5-Hydroxymethyldeoxycytidylic acid, 10–12
  synthesis of, 10–12
Hypochromicity as a test for DNA synthesis, 19

K

Kinase, deoxycytidylate, 15
Kinases, 12–17
  dGMP, 16
  deoxynucleoside diphosphate, 15–17
  thymidylate, in bacteria and HeLa cells, 173–174
    in liver, 171–172

L

Linkers, chromosome, 98, 108

M

Maps, genetic,
    *E. coli,* 146
    mutations, 248–249
    T-even phages, 131–135, 212
      in *rII* region, 213
Messenger RNA, *see* RNA
Microsomes, 292
Mitotic cycle, 155–156
  measurement of length, 156–157
Mitosis, synchronization, 181
Mutagenic specificity and the reaction of certain bases, 254
Mutagens, 217–218
  specificity of, 252
Mutations, 207–264
  classification of, 208–209
  forward, specificity of, 252
  induction, in TMV, 497–499
    kinetics of, 250
  phages for study of, 211–215
  reverse, 249
    specificity of, 253
  spontaneous, 245–248, 263–264
  suppressor, 254

N

Nitrogen-15 and DNA labeling, 78–79
Nitrogen mustard, *see* Alkylating agents
Nitrous acid, mutagenic effect of, 225–227
Nucleosides in relation to mitotic cycle, 159

P

Peptide bond
  formation of, 374–378
Peptide chain
  formation of, 373–374
Phage, *see* Bacteriophage
Phosphodiesterase, 23–24, 54
Phosphorus-32 as label for DNA, 18

Phosphorylase, polynucleotide, 280
Phosphorylation of deoxyribonucleotides, 12–16
Photodynamic effect of dyes, 225
Point mutation, 214
Polymerase, DNA,
  bacteriophage-specific, 54–56
  calf thymus, 50–54
  *E. coli*, 17–21, 40–45
Polymerase, RNA, 280–281
Polynucleotides, synthetic, 280
  use in coding studies, 282–285
Pools, metabolic, 311–314
Primer as a factor in controlling DNA replication, 181–183
Proflavin, mutagenic effect of, 224, 239
Protein synthesis, 353–400
  final stages of, 378–381
  kinetics of incorporation of label, 331–333
  requirement for DNA replication, 178–179
  sequence of bond formation, 374–378
Pyrophosphorolysis, 21

R

Radiation, indirect effects in mutation studies, 242
Radiations, ionizing, mutagenic effect of, 240–243
  ultraviolet light, mutagenic effect of, 243–245
Recombination, genetic, 82–88
  mechanism of, 87, 96, 107, 132, 137
  λ phage DNA, 86–87
Replication, *see* DNA replication
Revertants, 254
Ribonucleic acid (RNA)
  biosynthesis of, 272–286, 326–331, 335–341
  DNA-like, 273
  RNA, homologous to DNA, 275
  messenger, 275–276, 344, 382–386
    isolation of, 279
  soluble (*see also* transfer RNA), 368–369
  synthesis of, 271–286
    cell free systems, 397–400
    DNA-directed, 281, 386–390, 396–397
    relation to DNA synthesis, 186–188
  T2-specific, 274
  turnover of, 275–276
    upon phage infection, 278
Ribosomes, 292–350
  assembly of, 333
  effects of chloramphenicol, 298
    of magnesium on, 293–296
  fractionation of, 308–310
  nucleic acid components of, 305–306
  protein-RNA ratio, 300–302
  protein components, 302–303
  purification of, 299–302
  site of protein synthesis, 299
  structure of, 295, 306–307
  synthesis of, 321–350, 395–396
    site of, 346, 363–367
Ribotides, conversion to deoxyribotides, 169–170

S

Sedimentation analysis, *see* Gradient centrifugation
Serine as methyl donor, 7
Sulfur mustard, *see* Alkylating agents

T

Target theory, 241
Tetrahydrofolic acid, 7–9
Thalassemia, 450–453
  genetics of, 450
Thymidine-$H^3$, 74–76
Thymidine kinase, 13
Thymidylate kinase, 13
Thymidylic acid, 6–10
  synthesis of, 6–10
Tobacco mosaic virus (TMV) 477–517
TMV,
  amino acid composition, 493–495
  infectivity, 489–491
  protein, 484–489
    amino acid sequence, 486
  protein composition of mutants, 499–508
  ribonucleic acid, 487–489
  structure of, 479–483
  symptomatology, 491–493
  virulence and viability, 508–509

Transdeoxyribosidase, 4–5
Transfer RNA, 272, 356–362
  structure of, 359–360
Transformation, genetic, 133–134
  with lambda-dg DNA, 133–134
  with phage DNA broken by shear, 134–135
Transitions, 217, 256–257, 260–263
Transversions, 217
Triethylenemelamine as a mutagen, 233

U

Ultraviolet light, effects on pyrimidines, 245
  mutagenic effect of, 243

V

Variations, heritable, *see* Mutations
Viscosity measurements and tests for DNA synthesis, 18–19

[illegible] LABORATORY
LIBRARY